Theoretical modelling of reaction mechanisms of triazine and trinitrobenzene derivatives

Gooding, Stuart Robert

Award date:
2001

Awarding institution:
University of Bath

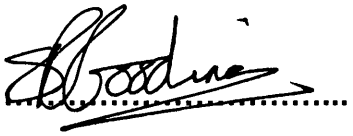
THEORETICAL MODELLING OF REACTION MECHANISMS OF TRIAZINE AND TRINITROBENZENE DERIVATIVES

Submitted by Stuart Robert Gooding

for the degree of PhD

of the University of Bath

2001



Stuart Robert Gooding

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THEORETICAL MODELLING OF REACTION MECHANISMS OF TRIAZINE AND TRINITROBENZENE DERIVATIVES

Table of contents

Page Number

Abstract

Fixation of reactive triazine dyes to cellulose is a topic of importance in the dyestuffs industry, where an issue of interest is the wastage of dyes due to competitive hydrolysis in the dyeing process. As part of trying to reduce these losses, it is desirable to obtain an understanding of the reaction mechanism. We have employed some methods of computational quantum chemistry to investigate this problem.

Calculations with semiempirical MO methods including continuum treatments of aqueous solvation show that nucleophilic aromatic substitution on derivatives of trinitrobenzene involves a σ-adduct (Meisenheimer complex) as the intermediate in a stepwise mechanism, in accord with experiment. However, similar calculations for derivatives of 1,3,5-triazine showed the σ-adduct to be a transition structure leading directly to a ion-molecule complex. It was not obvious whether this result was a genuine indication of a different mechanism for nucleophilic aromatic substitution of triazines or was an artefact of the method of calculation. A series of studies was therefore conducted to evaluate the reliability of both semiempirical and density-functional theoretical methods for modelling of reactions in solution.

Identity reactions of triazines with pyridine as nucleophile and leaving group were found to be stepwise, but contrary with experiment, so too were reactions with 4-nitrophenolate as nucleophile and leaving group. Calculations of gas-phase proton affinities for substituted pyridines and phenolates yields results in good agreement with experiment, whereas related calculations for pK_a values of pyridinium cations and phenols, using continuum solvation methods, are unsatisfactory. The problem is with the treatment of aqueous solvation.

DFT calculations with solvation predicted that ion-molecule complexes are indeed involved the addition of nucleophiles to chlorotriazine.

Acknowledgements

I wish to acknowledge the following people who have helped me over the past few years. Firstly I would like to thank my supervisor, Professor Ian Williams and industrial supervisor Dr David Buttar for assisting me through the various challenges encountered during my research. I would like to thank Dr David Buttar and Mr Phil Hoskin for access to their work carried out previously in this area. I would also like to thank the computational chemistry group especially Dr Stuart Firth-Clark, Dr Gail Rickard, Mr James Robinson, Dr Mark Russell and Dr Chris Rodriguez for their help and companionship throughout the course.

I would especially like to thank my parents for their assistance and support throughout my PhD. I would also like to thank them and my housemate, Mr Steve Rodway, for tolerating me throughout my write up period.

I would also like to thank Professor Ian Williams, Dr Mark Russell and Mr Steve Rodway and Miss Leonie Smith for proof reading my thesis and their constructive criticism of my thesis.

I would like to thank the Computational Chemistry Working Party (UK) for granting time allocations on Columbus at the Rutherford-Appleton Laboratory. I would also like to thank the support given to the purchase of the Silicon Graphics origin 2000 (Earth) on which most of my calculations were performed. I would like to thank the Electronics and Physical Sciences Research Council and Avecia limited for funding of this research.

Chapter 1
Introduction

Research into determining mechanistic pathways requires both the ingenuity of the researcher and the use of a various techniques to extract all the information about each particular mechanism. Calculations can be used both to guide and to complement experimental work, giving quantitative results for transition structures and intermediate complexes along a reaction pathway; this provides insight into the mechanistic nature of the reaction studied.

The focus of this work is the reactivity of 1,3,5-triazine units. These systems are important as they are used as dyes, especially for textiles and in inkjet printers[1-4]. Calculations provide a description of how the dye fixation process works at an atomic level, giving transition state and intermediate structures for the process.

1.1) A (very brief) history of dyes and dyestuffs

Dyes are a group of intensely coloured complex organic compounds that have been utilised to colour other textiles, paper, leather and many other materials. The craft of dyeing and the evolution of dyes several thousand years old and has progressed over the centuries mainly by empirical methods and serendipity. The focus of this research is on reactive dyes: dyes that have non-metallic groups capable of binding to the substrate.

William Henry Perkin discovered the first synthetic dye, Mauvein, in 1856 whilst searching for a cure for Malaria. This led to further research on dyes throughout the 19th century and the synthesis of azo dyes in the mid 1870's. The first series of fibre reactive dyes based on triazine, Procion dyes, were introduced by ICI in 1956. It was in the 1950's that the covalent chemical reaction of dye fixation was recognised and considered to be of technical importance.

The general characteristic structure of a reactive dye is shown below schematically in Figure 1.1; in some cases the reactive group is attached directly without a bridge.

Figure 1.1 – The general structure of a reactive dye.

The reactive group chosen (for example triazine) relies on the fact that the desired substrate (cellulose, protein and fibre groups) contain groups that are able to react as nucleophiles.

Dyes may bind to their target in three ways: mechanical retention, physical adsorption and chemical reaction. There is no clear distinction between physical and chemical dye-substrate interactions. However, the formation of covalent bonds by reactive dyes, as examined in this study, is definitely considered as a chemical process. Reactive dyes are coloured compounds which have groups that form covalent bonds between a carbon or phosphorus atom on the dye and an oxygen, nitrogen or sulphur atom of a hydroxyl (-OH), an amino ($-NH_2$) or a mercapto (-SH) group of the substrate. [5]

There are seven distinct classes of reaction undergone by reactive dyes, as follows. [5]

1) Groups that react by nucleophilic addition-elimination (heteroaromatic) substitution mechanism, for example triazinyl based dyes.
2) Groups that react by addition of the nucleophilic group of the substrate to a carbon-carbon double bond (C=C) on the reactive group.
3) Groups that react via several addition and elimination steps with a functional group HY.
4) Groups that react with the dye under acidic conditions.
5) Reactive dyes with phosphoric acid groups and phosphoric acid ester groups.
6) 'Reactive dyes' with thiosulphate ($-SSO_3H$) groups
7) Dyes with two reactive groups.

The dyes focused upon here, 1,3,5-triazinyl based dyes, fall into the first group and react by nucleophilic substitution. It is important to note that the reaction *in situ* involves competition between reaction with the desired substrate, for

example fibre, and that of hydrolysis. In order to produce a better leaving group catalytic quantities of diazabicyclo-[2,2,2]-octane (DABCO) or 1,1-dimethyltriazine were added to improve the reaction. [6] However, this practice has ceased as it decreases the ratio of reaction with fibres to that of hydrolysis. This competition between the dye fixation and hydrolysis (as mentioned previously) is of industrial importance. The maximisation of dye fixation over hydrolysis prompted this investigation into the mechanism of fixation.

1.2) Dye Fixation

The fixation of dyestuffs and reactive dyes in particular to a recording medium, for example paper or cotton, can be achieved by nucleophilic aromatic substitution under alkaline conditions. Reactive dyes include those based upon the 1,3,5-triazines. The structures of two industrially important dyes are shown in Figures 1.2 and 1.3, each of which contains both a chromophore and leaving group attached the triazine ring.

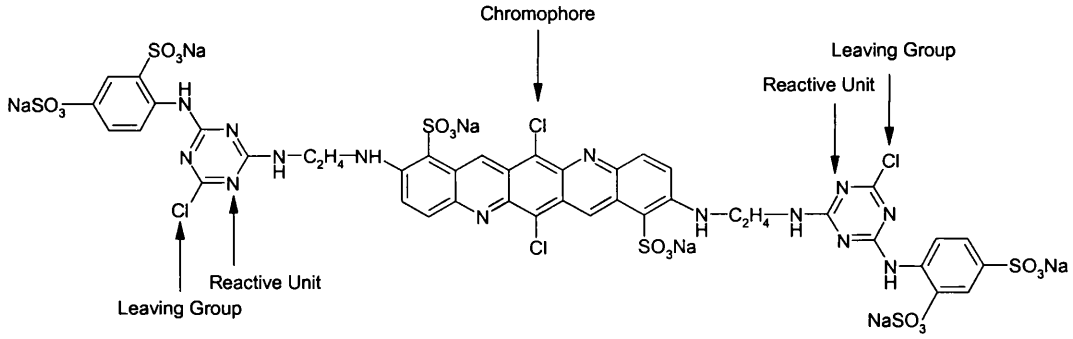


Figure 1.2 – Procion Blue H-EGN – A popular textile dye

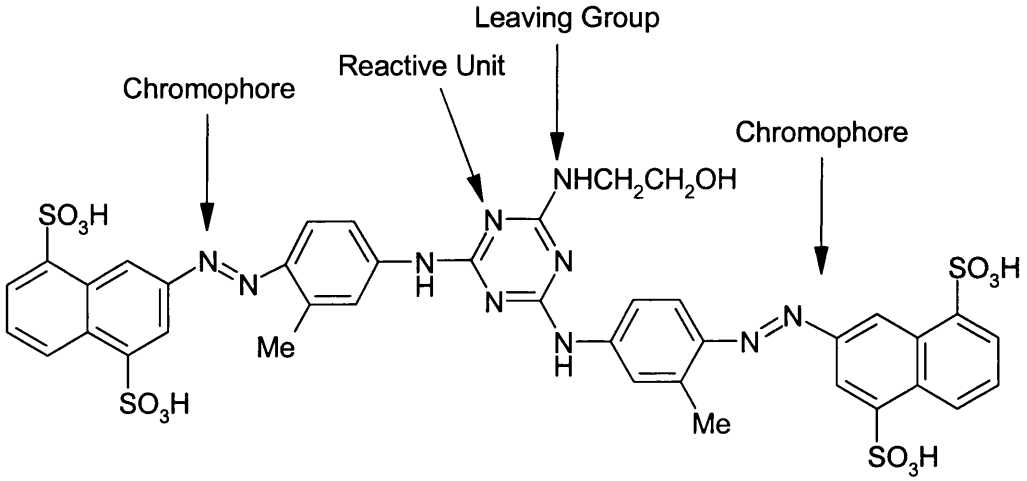


Figure 1.3 – CI Direct Yellow 86[4] – An inkjet printer dye

The overall reaction for binding a triazinyl based dye to a cellulose fibre of cotton or paper is shown schematically in Figure 1.4. A hydroxyl group of cellulose is thought to be ionised, under alkaline conditions, and to act as a nucleophile in a substitution process that forms a covalent bond to the nucleophilic group and causes the leaving group to depart from the triazine.

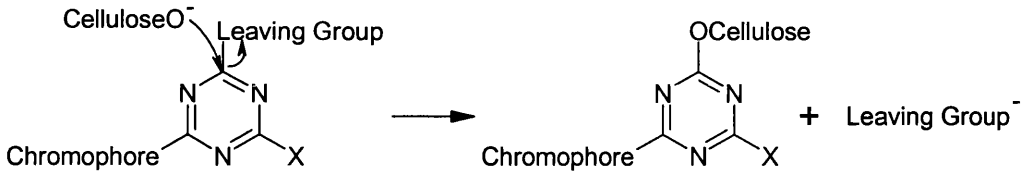


Figure 1.4 – Schematic diagram of fixation

It is important to achieve selectivity in favour of the cellulose derived alkoxy fixation ($^-$OCellulose) over hydrolysis ($^-$OH). If the dye is readily hydrolysed then a lot of it is wasted in the dyeing process. The efficiency of fixation is affected by both the recording medium and the dye used. In order to optimise the efficiency of fixation, it is important to understand how the fixation mechanism works at a molecular level.

1.2.1 Stepwise Mechanism (A_N + D_N) and Meisenheimer reaction

If an appropriate leaving group is present on an activated ring system, nucleophilic aromatic substitution can, in principle, take place by one of several mechanisms. Systems involving strong electron-withdrawing groups, for example trinitrobenzene or triazine, react by a stepwise addition/elimination (A_N + D_N) mechanism, as shown in Figures 1.5 and 1.6. This mechanism is rare in systems with benzene derivatives except when activated substrates are used, for example nitro-substituted compounds. In this mechanism the nucleophiles forms a bond with the substrate, giving a σ-adduct as an intermediate, then the leaving group departs. The first intermediate isolated by Meisenheimer[7] in 1902 involved the reaction between ethyl picrate and methoxide ion, shown in Figure 1.5. Intermediates of this type are stable salts, referred to as Meisenheimer or Meisenheimer-Jackson salts. [8]The existence and structure of these salts has been proved experimentally by various methods, [9-14]for example, NMR, [15-19]and x-ray crystallography. [20, 21]The Meisenheimer reaction has been widely

studied for activated benzene derivatives. However, at present, there have not been extensive studies of triazine systems.

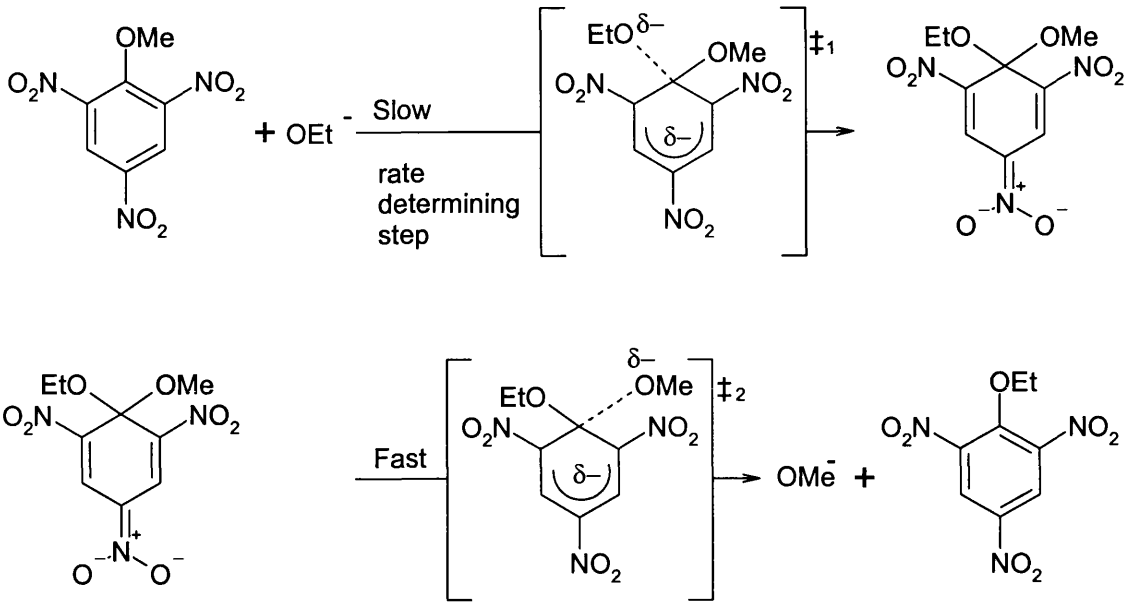


Figure 1.5 – The stepwise (A_N + D_N) Meisenheimer reaction mechanism[7] of ethyl picrate with a methoxide ion.

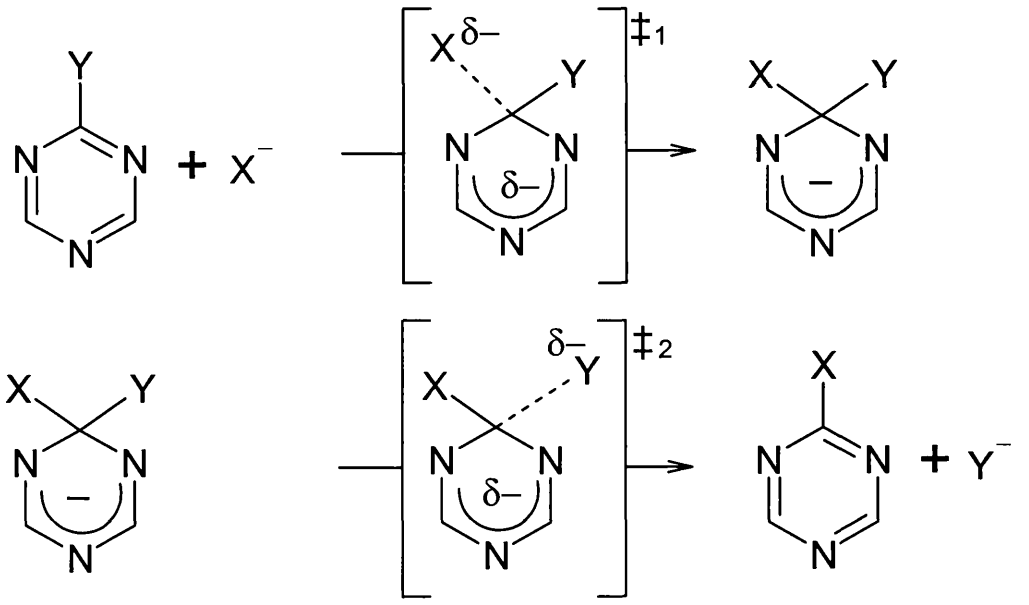


Figure 1.6 – The proposed mechanism for the stepwise (A_N + D_N) reaction for a triazine type compounds.

An alternative stepwise mechanism is the (D_N + A_N) reaction, shown in Figure 1.7; an appropriate leaving group for this mechanism would be for example, the azo group (-N≡N^+).

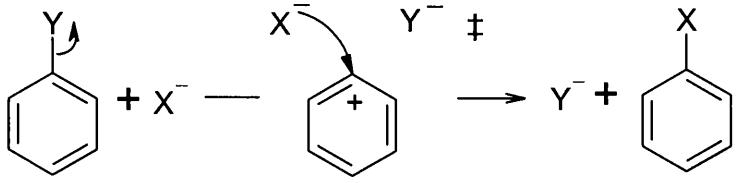


Figure 1.7 – The mechanism for the stepwise (D_N + A_N) reaction.

1.2.2) Concerted Mechanism (A_ND_N)

An alternative mechanism for nucleophilic aromatic substitution is the concerted (A_ND_N) mechanism, shown in Figure 1.8. It is possible that this mechanism is prevalent in the reactions of triazines.

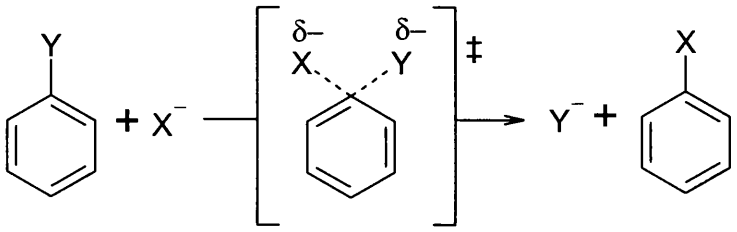


Figure 1.8 - The proposed mechanism for the concerted (A_ND_N) reaction for triazine based compounds.

One such example is the reaction of 4-nitrophenolate anions with 4,6-dimethoxy-1,3,5-triazine published by A. Williams and co-workers, [22]shown in Figure 1.9.

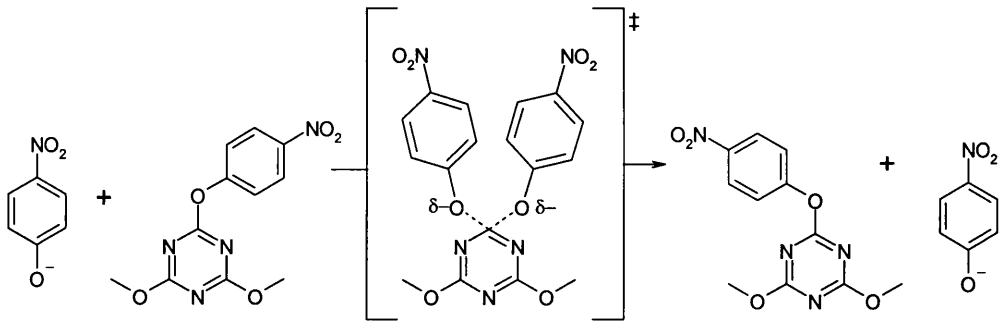


Figure 1.9 – The concerted (A_ND_N) mechanism for the reaction of 4-nitrophenolate anions with 4,6-dimethoxy-1,3,5-triazine.

There are several other mechanisms for nucleophilic aromatic substitution of benzene derivatives, involving electron-transfer chain processes. These include a radical initiated reaction or oxidatively initiated nucleophilic bimolecular substitution that can be induced by a variety of methods, for example electrochemically.

Chapter 2

Introduction to computational methods

When modelling a chemical reaction, a variety of computational techniques can be used. These techniques can be used to explain complex chemical behaviour and assist chemists in the understanding of structure and energy variations throughout a reaction. Such studies can be used to assist practical chemists in optimising the conditions for the desired results, for example to optimise speed or selectivity. The accuracy of the computational methods used dictates the confidence that can be given to the results; generally a technique is more accurate the less approximations used, and therefore the more computationally expensive it is. It is therefore important to use an appropriate level of theory for the system studied as they vary widely in both accuracy and computational expense.

When modelling a chemical reaction, an explicit representation of the interacting electrons is required. This necessitates the use of Quantum Mechanical (QM) techniques, [23-29]except in special cases where highly parameterised models are used. QM techniques explicitly represent electrons; this gives the potential to describe systems accurately and as such it is possible investigate properties that depend on the electronic distribution. It is therefore possible to model the behaviour of electrons during a reaction. Ideally, QM would be used for all cases studied but it is computationally expensive, depending on the approximations applied. Even using these approximations only moderate sized molecules can be modelled by fully QM calculations due to computational constraints.

In general, the size of the systems investigated in this study are too large for high level *ab initio* techniques; for example, Møller-Plesset perturbation theory[30] calculations. Therefore lower level *ab initio* and density functional methods as well as semiempirical methods have been applied. In an *ab initio* approach, all the integrals describing the system are treated in full for the specific basis set used; this may or may not include electron correlation effects. Density functional methods model the behaviour of the whole system as a function of electron density; this method includes electron correlation effects. Semiempirical molecular orbital methods provide an alternative approach where all the electrons

(except the valence electrons) are represented by a core function. The core function eliminates the need for calculation of many of the computationally intensive integrals, as they are parameterised in accordance with empirically derived data or neglected completely. This considerably reduces the amount of computational expense involved.

2.1) A (brief) summary of the basics of quantum mechanics

The basis of Quantum Mechanics (QM) can be represented as a series of postulates. These encompass the behaviour of all known forms of matter; this of course, includes the molecules, atoms and electrons investigated in this study It is impossible to encompass with great depth the whole range of QM techniques in this Section; therefore, I wish to give a brief overview of QM. Much greater detail is contained in an abundance of texts published on this subject. [23-29]

The rapid development of computer hardware and software allows ever larger systems to be studied by *ab initio* and Density Functional Theory (DFT) techniques and the use of semiempirical methods extend the range of QM calculations possible. These techniques, depending on the type of calculation being performed, are at present employed for systems with up to 40 atoms, and semiempirical techniques are used for systems with up to 100 atoms.

The central equation in QM is the Schrödinger equation, shown here in its time independent form: -

$$\hat{H}\psi_n = E_n\psi_n \qquad (2.1)$$

where $\hat{H}$ is the Hamiltonian operator that operates on the wavefunction ψ_n corresponding to this state n, and E_n is the energy of state n. The total wavefunction ψ_n of a molecule is a mathematical function constructed from a set of atom centred functions that describes the properties of the system. Once a wavefunction is known, any observable can be calculated by applying an operator to the wavefunction in such a way that they result in a product of the wavefunction and a set of allowed values (eigenvalues) of the property. There are a range of operators that can be applied to the wavefunction, for example the Hamiltonian operator, $\hat{H}$, which yields the systems energy. The calculation of an

observable results in an average value from applying an operator on the wavefunction. The energy can be determined from the following equation: -

$$E(observable) = \frac{\int \psi^* \hat{H}(operator) \psi d\tau}{\int \psi^* \psi d\tau} \qquad (2.2)$$

Where $d\tau$ indicates integration over all space and ψ^* is the complex conjugate of the wavefunction, ψ. The Hamiltonian operator consists of a sum of kinetic and potential parts, which include terms related to the electron kinetic energy, electron-nuclear attraction, electron-electron repulsion and the nuclear-nuclear repulsion respectively.

The wavefunctions that result from the Schrödinger equation for the hydrogen atom are referred to as orbitals. In polyatomic molecules, the orbital approximation is used which treats each electron as consisting of a hydrogen-like atomic orbital (AO). In the case of Boron, the electron configuration is $1s^2 2s^2 2p^1$. This is described as having separate functions associated with two electrons being in the 1s AO, two more being in the 2s AO and one being in the 2p AO. In general, the total wavefunction, ψ, can be considered as consisting of a number of one-electron atomic wavefunctions, χ_i.

There are complications when solving the Schrödinger equation for atoms with three or more particles as the Schrödinger equation cannot be solved exactly due to the three-body effect and the effects of electron spin. Consequently, all calculations with three or more particles can only be solved approximately. The Pauli exclusion principle states that no two electrons can be represented by the same four quantum numbers. Therefore no two electrons in the same orbital can have the same spin. Consequently, each of the one-electron functions has to be multiplied by a spin factor. The electron configuration for boron for example can be written in a shorthand form as $1s^{\alpha} 1s^{\beta} 2s^{\alpha} 2s^{\beta} 2p^{\alpha}$, where α and β represent the two opposite directions of electron spin.

The Schrödinger equation can be solved for systems larger than two particles if the motion of the electrons and nuclei are decoupled using the Born-Oppenheimer approximation. In this approximation, the positions of the nuclei are fixed whilst the electrons are allowed to move. This works well, as the mass

of the nuclei is considerably larger than that of an electron, therefore the electrons move almost instantaneously compared to the nuclei.

In principal molecular orbitals can be treated in the same way as atomic orbitals, the molecular orbital approximation can be assumed; thus, a molecular orbital, Ψ, is given by adding together the effects of all of the individual electrons within the molecule. In most QM packages the molecular orbital is formed as a linear combination of basis functions, as described in Section 2.1.2.

2.1.1) Self consistent field

Generally, the Schrödinger equation central to QM is solved using a self-consistent field[31] (SCF) method. This method assumes that any one electron moves in a spherically average potential caused by the other electrons and nuclei. In order to get a spherically averaged potential the wavefunctions of the other particles are assumed to be known. As these are generally not known, an approximate form of the wavefunction is used. The Schrödinger equation is then solved for one electron and the procedure is repeated for all the other electrons in the atom. This gives improved wavefunctions for all the electrons. The process is then iterated using the improved set of wavefunctions until convergence. The self-consistent wavefunctions produced using this method are a good approximation to the true wavefunction.

2.1.2) Basis sets

The set of atomic orbital functions in an linear combination of atomic orbitals – molecular orbital is called a basis set. When using *ab intio* and DFT methods, it is important to use an appropriate basis set. An infinite set of basis functions must be used to represent orbitals exactly. The result of such an equation would be at the HF limit. However, this would not be the exact ground-state energy, as HF neglects electron correlation. As an infinite series of basis functions is computationally intractable, in practice the basis sets are truncated. This leads to an error called "basis set truncation error". Therefore, there has to be a balance between precision and computational expense. However, keeping the number of basis functions low and choosing the basis functions carefully can produce an efficient result.

There is a variety of basis sets used and the most common ones have systematic shorthand names. These basis sets usually include all the AOs in the shell. The most common minimal basis sets are Slater-type orbitals[32], for example is STO-3G. In general, these are written as STO-nG, where n is the number of Gaussian[33, 34] functions used to represent each Slater-type orbital. In practice, at least three Gaussian functions are needed to describe the AO properly. Therefore the STO-3G basis set is generally considered the absolute minimum in *ab initio* calculations.

These minimal basis sets give remarkably good results although this is mainly due to a fortuitous cancellation of errors. There are several problems with minimal basis sets. These problems are encountered with atoms with nearly full valence shells. This is because the minimal basis sets use the same number of functions for the valence region independent of the number of electrons present in it. Other drawbacks encountered with this approach are that the sizes of the atoms are fixed and therefore the atoms do not fully describe the non-spherical electron distribution properly. These effects combine so the description does not adapt correctly to the surrounding environment and as such are not transferable. These problems can be overcome by using more than one basis function for each orbital.

A basis set with double the minimal basis set functions is called a 'double zeta' basis and describes the atoms with improved accuracy. Likewise a 'triple zeta' basis set has three times the minimal basis sets functions this provides a further increase in accuracy. These are made of a linear combination of contracted functions that give an improved overall result. An alternative approach is to double the number of functions in the valence shells whilst keeping single functions for the inner shells. The logic behind this method is that the valence electrons affect the atomic properties more than the core ones, and hence they need to be modelled more accurately. The notation used for such split valence basis sets can be described by examining the 6-31G basis set. This uses six Gaussians to describe the core orbitals; whilst the valence orbitals are described by four Gaussian functions, three contracted parts for the inner region and one for the outer region. The most commonly used split valence basis sets are 3-21G, 6-31G and 6-311G.

These split valence basis sets neglect the effect of polarisation on the system, so polarisation functions are added to produce a more precise representation. The polarisation functions are indicated by either an asterisk, for example 6-31G* or by a suffix of the type of polarisation function added, for example 6-31G(d,p). In the first nomenclature, one asterisk indicates the inclusion of the polarisation effects for heavy atoms (i.e. not hydrogen atoms) and two asterisks indicates the inclusion of polarisation on all atoms. In the second nomenclature, the (d,p) in 6-31G(d,p) indicates one extra d-type orbital on each heavy atom and one extra p-type orbital on the light (hydrogen) atoms.

A further refinement to improve the flexibility of the basis set, especially on anionic species, is the addition of diffuse functions. These functions give the charge the ability to diffuse further away from the nucleus. This assists the representation as the electron density away from the nucleus is generally underestimated when using Gaussian type orbitals. This is especially the case in the case of negatively charged species. These diffuse functions are denoted by the addition of a plus, for example 6-31++G(d,p), where one '+' indicates the addition of a single diffuse function to all heavy atoms and two '++' indicates the inclusion of an extra diffuse function on all the atoms including the hydrogens. The addition of polarisation and diffuse functions provide a much-improved molecular representation, especially if there is a highly uneven distribution of charge.

2.1.3) Hartree-Fock Theory

As described earlier, the true wavefunction cannot be generated for anything other than the simplest system; so there needs to be some indication of the quality of a proposed wavefunction. This is given by the *variation theorem*, which effectively states that the approximation to the energy of a wavefunction is always more than the true wavefunction. Consequently, the best approximation has the least energy. At an energy minimum, the first derivative of the energy is equal to zero. This is the basis of the Hartree-Fock (HF) equations, subject to the condition that the Molecular Orbitals (MO) remains orthonormal. MOs are said to be orthonormal when the set of functions describing them are both normalised (where the integral of the function over all space is unity) and orthogonal (where

the overlap integral of the functions is zero). In the HF approximation the correlation of motion between individual electrons is treated in an average way, and therefore their mutual repulsion is disregarded. In HF theory the wavefunction is calculated as a single determinant; that is to say if the electrons are interchanged, the sign of the wavefunction is reversed. This is to satisfy the requirements for anti symmetry outlined in Pauli's exclusion principle. The HF technique can be used to produce a good representation of the distribution of charge throughout the system; however the correlation effects are neglected. An extension of the HF theory, configuration interaction, which includes the correlation effects is described in Section 2.1.6, this gives a very good approximation of the true wavefunction.

The HF equation for orbital $\phi_a(1)$, where an electron 1 is assigned to orbital ϕ_a, is

$$f_1\phi_a(1) = \varepsilon_a\phi_a(1) \tag{2.3}$$

where ε_a is the orbital energy and f_1 is the Fock operator: -

$$f_1 = h_1 + \sum_u \{J_u(1) - K_u(1)\} \tag{2.4}$$

Here h_1 is the core Hamiltonian for electron 1, u represents all the spin orbitals, J_u is the Coulomb operator and K_u is the exchange operator. The Fock operator is an effective one-electron energy operator, describing the kinetic energy of an electron, the attraction to all the nuclei and the repulsion to all the other electrons (via the J and K operators). The Coulomb operator takes into account the repulsion between electrons and the exchange operator represents the effects of spin correlation. Each orbital has to be calculated for the corresponding Fock operator, which in turn depends on all of the other electrons present. To solve a HF equation it appears you need to know the answer first; this of course this is impossible. Therefore in practice a self-consistent field (SCF) approach has to be used, as described in Section 2.1.1. There is an infinite series of eigenfunctions for the Fock operator; however, in practice only a finite number of eigenfunctions can be used. These are arranged by increasing orbital energy ranging from the lowest occupied orbitals through to unoccupied orbitals, or virtual orbitals.

Practically, these are generally calculated using matrices related to the Hamiltonian using the Roothaan-Hall equations. When solving the Roothaan-Hall equations they can be expressed in terms of a matrix equation, where the Fock matrix is a K x K square matrix: -

$$FC = SCE \qquad (2.5)$$

where F is the Fock operator, C is the elements of the K x K matrix, S is the overlap matrix and E is the diagonal matrix whose elements are orbital energies. One common approximation of semiempirical methods is that the overlap matrix is set equal to the identity matrix, although this does not mean that the overlap integrals are set to zero when calculating the Fock matrix. Thus the Rotthaan-Hall equations are simplified: -

$$FC = CE \qquad (2.6)$$

2.1.4) Density Functional Theory

Density functional theory[35, 36] (DFT) is an alternative, accurate approach for describing the system to the *ab initio* methods already described. DFT is based on the Hohenberg-Kohn theorem, which states that the ground-state properties of a system are dependent on the electron density, ρ. The main advantage of the DFT approach is that, unlike HF *ab initio* methods, it includes electron correlation. This is achieved at a reasonable computational cost compared to high level *ab initio* techniques such as Møller-Plesset perturbation theory; [30, 37]consequently larger systems can be studied accurately. The results of DFT for d-block metals generally show a better correlation with experimental results than HF methods.

The total ground-state electronic energy, E, can be written as a function of the electron density, ρ. This can be considered as being made up of a series of terms related to the kinetic energy, $E_{KE}(\rho)$, the electron-nuclear interaction energy, $E_C(\rho)$, electron-electron interaction energy, $E_H(\rho)$, and $E_{XC}(\rho)$ which represents the exchange and correlation interaction terms. Thus the total electronic energy can be expressed as a function of the electron density: -

$$E(\rho)=E_{KE}(\rho)+E_C(\rho)+E_H(\rho)+E_{XC}(\rho) \qquad (2.7)$$

The exchange-correlation term, $E_{XC}(\rho)$, is the only term which can not at present be calculated exactly and therefore this is the largest source of error in DFT. An approximation needs to be used, the most common one being the local density approximation. This is based on the behaviour of a uniform electron gas where the electron density is constant throughout the system. When using the local density approximation the exchange-correlation energy can be written: -

$$E_{XC}(\rho) \cong \int \rho(r) \varepsilon_{XC}[\rho(r)] dr \qquad (2.8)$$

where $\varepsilon_{XC}[\rho(r)]$ is the electron-correlation energy per electron in a constant density electron gas. The exchange-correlation term is refined to take into consideration the change in the electron density using a non-local correction. This usually involves a function of the gradient of the electron density, $\nabla\rho$.

2.1.5) Semiempirical techniques

The use of *ab initio* and density functional theory techniques[35, 36] is computationally expensive for larger systems. The majority of the time taken to calculate a molecule using *ab initio* QM methods is spent computing and manipulating integrals. In order to allow large systems to be treated, semiempirical approaches[29] may be used in which these integrals are generally approximated or neglected (as described later), thus speeding up the calculation immensely.

One of the main approximations used to simplify the calculations in semiempirical methods, is the neglect of the explicit representation of the core electrons. The effect of these core electrons is included in the representation of the nuclear core. The valence electrons are calculated explicitly as they are the electrons that are most associated with chemical interactions. Therefore, these valence electrons are therefore the most important electrons to be investigated. The semiempirical approaches generally use basis sets that consist of Slater-type orbitals, which are orthogonal; this allows more approximations.

Semiempirical methods are based on a parameterised adaptation of the Hartree-Fock Hamiltonian. It is important to note that even in the simplest methods some of the overlaps are considered. However most semiempirical techniques involve a simplification of the matrices used to calculate orbital overlap. There is a range

of semiempirical techniques which vary in the way this overlap is treated. These techniques include zero-differential overlap (ZDO), complete neglect of differential overlap (CNDO), intermediate neglect of differential overlap (INDO) for one-centre integrals, and neglect of diatomic differential overlap (NDDO). These approaches are rarely used today, due to the rapid improvement in computer power since they were introduced and the extremes of the approximations used. However, techniques derived from these approaches, notably by Dewar and colleagues, [38-44]are in use today. The earlier of these are the modified INDO (MINDO/3) method[41-44] that although similar in theory to INDO, includes a lot more parameterisation from experimental data. The MINDO/3 approach had some significant limitations such as errors in the calculation of the bond angles, and the heats of formation of unsaturated molecules to become too positive. Another approach was then derived called the modified neglect of diatomic overlap (MNDO) method[38, 39] based on the NDDO method. This approach had significant advantages over the MNDO/3 approach, for example, the use of monatomic parameters allowed the inclusion of different core-core repulsion terms. The range of elements was also increased, especially when the approach was modified further to include *d*-orbitals. Nonetheless, a range of errors are still inherent in the technique.

The two most commonly used semiempirical methods today are the Austin Model 1 (AM1) [40]and Parameterised Model 3 (PM3) [45]approaches. The AM1 and PM3 models are both based on the same Hamiltonian; however they have different parameterisation systems. These models eliminate many of the difficulties with the original MNDO method by modifying the core-core term using spherical Gaussian functions. The AM1 and PM3 techniques use both attractive and repulsive Gaussian functions to counter the overestimate of repulsion between nuclei in MNDO.

The number of parameters used in AM1 is much greater than MNDO with the number of parameters per atom increased from 7 in MNDO to between 13 and 16 in AM1. The AM1 parameters were derived from empirical results and chemical intuition. The PM3 parameters were derived using an automated parameterisation procedure devised by J.J.P. Stewart. [45]The AM1 and PM3 techniques both give approximately the same level of accuracy.

2.1.6) Electron correlation

When the SCF method is used, the electrons are assumed to be moving in an average potential and the influence of neighbouring electrons are not treated correctly. The motion of the electrons is correlated so the electrons present can avoid each other, thus lowering the energy of the system. This is included implicitly in the use of the empirical parameterisation in semiempirical techniques, however it is not included in some *ab initio* calculations, notably in Hartree and Hartree-Fock theories. There are various ways of including correlation effects in an *ab initio* calculation. One of the most common methods is called configuration interaction (CI) where the electronic state contains a representation of the excited configurations. This is achieved by a linear mixing of the wavefunction, in which the different configurations are generated by excitation of electrons into virtual orbitals. Therefore, a general form of a CI wavefunction can be written: -

$$\Psi = c_0\psi_0 + c_1\psi_1 + c_2\psi_2 + c_3\psi_3 + c_4\psi_4 + c_5\psi_5 + \ldots \qquad (2.10)$$

where ψ_0 is the single-determinant wavefunction that is found from the SCF HF approach, the other wavefunctions, ψ_1, ψ_2, ψ_3 ... depict the configurations, say where the electrons have been promoted to virtual orbitals. The coefficients c_0, c_1, c_2, c_3 ... are selected to produce the minimum energy of the system in accordance with the Variational Principle.

When using CI, each extra term representing another excited configuration provides a better description of the true wavefunction. The full CI gives the best representation of the wavefunction for the chosen basis set. There is a balance encountered between precision and computational time. Consequently, the series is generally truncated and various approximations are used. One common approximation is the frozen core approximation where the orbitals corresponding to the core electrons are neglected. Some excited states do not mix with the ground state and therefore do not reduce the overall energy. This is described by Brillouin's theorem, which states that the matrix elements between the ground state and single excitation state are zero. However as single excitations interact with other excited states they can be included.

The CI energies for differently sized systems are incomparable. This size-consistency defect may be addressed using various methods. The most common method is Møller-Plesset perturbation theory[30, 37] this is expressed as a power series summing a solvable 'zeroth-order' Hamiltonian, H_0, and a perturbation, V_n, where *n* is the order of power: -

$$H = H_0 + \sum_n \lambda^n V_n \qquad (2.11)$$

where λ^n is a parameter with a value between zero and one; when λ is zero the Hamiltonian is equal to the zeroth order Hamiltonian, whereas when λ is one then the Hamiltonian equals its true value. The most popular level of Møller-Plesset theory used is second order Møller-Plesset (MP2) perturbation theory; this is the first level of theory that gives an improvement over HF theory. As at the MP2 level of theory the calculations scale badly and therefore this level of theory is generally restricted to treating compounds of less than 20 atoms. Third (MP3) and fourth (MP4) order Møller-Plesset methods are also incorporated into most *ab initio* packages, for example Gaussian98. [46]

2.2) Solvation effects

Most chemical reactions take place in solvent, therefore it is important to model the effect of solvation on a system. The behaviour of a system can be completely different in the presence of solvent; the solvent acts to perturb the gas phase system. In some reactions solvent molecules are directly involved in the reaction studied. In such cases, the solvent molecules are an integral component of the system and have to be treated explicitly as if they were part of the solute. There are various techniques of modelling the effect of solvation. These techniques range in both accuracy and computational expense from 'bulk medium' continuum models through to discrete representation of the solvent molecules. There is a large amount of literature published that covers the solvation treatment area[29, 47-49] therefore only an overview of the field is given here. Dielectric and conductor-like continuum models have been mainly used in this study, consequently these methods will be focused upon here also only a brief overview of an explicit treatment of solvation such as a 'supermolecule' approach will be given.

The dye fixation systems primarily investigated in this study all occur in solution therefore it is important to include the effect of solution when modelling them. This can be achieved using an explicit, partially explicit or continuum approach. The use of an explicit solvation requires sampling of phase space, as described in Section 2.4; this sampling of the different solvent orientations, which can influence a reaction, is computationally expensive. One reason for this computational expense when explicitly representing a solvent is the increase in the number of degrees of freedom that need to be considered. If 200 molecules of water are treated explicitly 1800 degrees of freedom need to be considered, 9000 degrees of freedom for ethyl ether and 16,200 degrees of freedom for 1-octanol. [47]Condensed phase systems generally have a large amount of local energy minima. This complicates the finding of transition states and following of reaction paths.

The use of continuum methods avoids the need for an extensive sampling of phase space of the potential energy surface. The averaging of the properties of the implicit molecules is reflected in the properties attributed to the bulk continuum system. There is a wide range of continuum models that have been presented for use with particular theoretical treatment of the solute. The most common continuum models focus on embedding a solute into a dielectric continuum of a certain permittivity, ε, to model the electrostatic effect of the solvent. The solute forms a 'cavity' within the dielectric; this cavity is usually the shape based upon the van der Waals surface of the solute. The choice of where the solute 'ends' and the continuum 'begins', that is to say the boundary, and the shape of the cavity is indistinct as there is no real physical quantifiable boundary. In some older continuum methods the shape of the solute cavity within the continuum was assumed to be either a sphere or and ellipsoid. The oldest models are based on multipole expansions that can be used to implement further approximations. In some cases the multipole expansion used was even truncated to the dipole level. These are extreme approximations and with present computational constraints, are unnecessary assumptions except in extreme cases. Other terms that are more commonly included (due to the need for more precision in modern models) are those related to the van der Waals and solvent cavity formation energies.

Continuum solvation models originally were designed for describing simple bulk physical effects, for example electrostatic effects. This relies upon expressions especially those derived by Born[50] and then Onsager; [51]these describe a solute as a point charge/dipole respectively within a medium of dielectric constant, ε. The work done to charge the point charge/dipoles is equivalent to the solvent's electrostatic contribution to the free energy. For a charged spherical system this work done can be calculated from $\frac{q^2}{2\varepsilon a}$ where q is the point charge and a is the radius of the cavity. Therefore, as the permittivity of a vacuum is unity, the difference between charging the ion *in aquo* compared to *in vacuo* is $-\frac{q^2}{2a}\left(1-\frac{1}{\varepsilon}\right)$ in atomic units. The total free energy of a system is the sum of the Coulombic energy and the Born free energy; this is for a *N*-atomic system with particles of radii, a_i, and charges, q_i: -

$$G_{elec}=\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{\varepsilon r_{ij}}-\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{2.12}$$

The Coulomb interaction can then be rewritten in terms of the sum of its *in* vacuo and *in aquo* components. The difference between the Coulombic energy from the charge-charge interactions *in vacuo* and the Coulombic energy *in aquo*, ΔG_{elec}, is this is given by the generalised Born equation: -

$$\Delta G_{elec}=\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{q_i q_j}{\varepsilon r_{ij}}-\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{2.13}$$

The generalised Born equation can be rewritten to contain a single term this has the form: -

$$\Delta G_{elec}=-\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{f\left(r_{ij},a_{ij}\right)} \tag{2.14}$$

where $f\left(r_{ij},a_{ij}\right)$ is a function of the interparticle distances, r_{ij}, and the Born radii, a_{ij}.

A separate major set of methods is based on solving the Poisson equation for the interface between the solvent and solute. The Poisson equation relates the variation in the potential, ϕ, within a uniform dielectric to the charge density, ρ, for a location, r. In atomic units this is:

$$\left(\left(\frac{\partial^2}{\partial x^2}\right)+\left(\frac{\partial^2}{\partial y^2}\right)+\left(\frac{\partial^2}{\partial z^2}\right)\right)^2\phi(r)=-\frac{\rho(r)}{\varepsilon_0\varepsilon} \tag{2.15}$$

The programs used include various dielectric continuum models. The MOPAC 93 program contains the COSMO continuum[52] and AMSOL 6.1[53] contains a variety of solvation models, the latest of which is solvation model 5.4 (SM5.4*x*). [54-56]The Gaussian 98 program[46] has a variety of continuum models, a solely dielectric Polarised Continuum Model (PCM) [57, 58]and the conductor-like screening model COSMO. These methods both treat the electrostatics of the solvent in a similar way using either by methods based upon the generalised Born approximation. The COSMO Polarised Conductor Model (CPCM) [59]is based upon a grid-wise solving of an approximation to the Poisson equation. This represents the electric polarisation of a dielectric medium by a polarised charge density at the solute/solvent boundary. The polarisation charge of the solvent then polarises the solute self-consistently by numerically solving the Poisson equation for the solute/solvent interface. This is computationally expensive to calculate, especially during an optimisation if the cavity is optimised, where there is a significantly change in cavity shape.

In common with all of the existing dielectric continuum models, the dielectric is treated as a homogenous bulk; although there has been discussion about the implementation of an inhomogeneous dielectric continuum. [60-63]The techniques also both assume the boundary of the solvent, continuum, and the solute by taking account of the molecular shape of the solute studied. These techniques both treat the solute as a fixed size set of atom centred spheres along a non-fluctuating distinct solvent/solute border. The techniques differ in the treatment of the solvent. The SM*x* model[47, 54-56, 64-68] treats the solute with polarised charges derived from either a rank four charge model or using an approximate QM method including correlation interaction. The COSMO model uses the polarised charges obtained from semiempirical calculations (these include correlation interaction implicitly through parameterisation). The models also differ as the electronic terms in SM*x* unlike COSMO includes non-electrostatic contributions. The SM*x* models include empirical elements that compensate for the approximations used in the calculation of the electronics and various other effects. The entire range of continuum models, for example COSMO and SM*x*, all contain approximations and thus the use of an appropriate model is application dependent and subject to their availability in common software packages.

2.3) Molecular Mechanics

Due to the computational expense it is impossible to study large molecules using rigorous quantum mechanical methods. When investigating these systems molecular mechanics (MM) [27-29]is usually used. One of the goals of molecular modelling is to simulate the three-dimensional properties and structure of the molecule being investigated. MM techniques use various approximations to simplify the calculations involved in this process. The algorithms used produce a mathematical model of the properties of a system. In MM these algorithms are based on the laws of classical physics. As MM is not used in this study only a brief overview will be given.

MM works by using an empirical fit to calculate, for instance, the potential energy surface of the molecule, thus creating a forcefield. The two important features of a forcefield are its simplicity and its transferability throughout a variety of systems. The parameters used for small molecules should be easily transferred to much larger molecules such as proteins. The accuracy of the forcefield produced is dependent on the quality of the potential functions and parameters used. There are several disadvantages to this model compared with a quantum mechanical model, the main one being that it does not use an explicit representation of the electrons.

In the MM representation, the molecules are represented by a collection of soft spheres joined by 'springs'. The strength of these springs differs depending on the various bond strengths. The motion is determined by the laws of classical physics and is represented by a combination of the intra- and inter-molecular forces acting on the system. When characterising an atom using a forcefield, extra data is required to describe the atom types in the molecule. This information is usually related to the hybridisation state and the environment in which the atom is located.

The forcefield is calculated from the functions representing the bonded (through bond) and non-bonding (through space) interactions. The total potential energy, E_{total}, is calculated as follows: -

$$E_{total} = E_{bond\ length} + E_{bond\ angle} + E_{torsion} + E_{non\text{-}bonded} \quad (2.16)$$

where the energy is considered as consisting of the bond length stretching energy, $E_{bond\ length}$, bond angle bending energy, $E_{bond\ angle}$, and torsion angle energy, $E_{torsion}$, along with non-bonded interactions, $E_{non\text{-}bonded}$.

One general four-term functional form of the algorithm used in a molecular mechanics force field is as follows: -

$$E_{total} = \sum_{bonds} \frac{k_i}{2}(l_i - l_0)^2 + \sum_{angles} \frac{k_i}{2}(\theta_i - \theta_0)^2 \quad (2.17)$$

$$+ \frac{v_n}{2}(1 + \cos(n\omega - \gamma))^2$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N-1} \left(4\varepsilon_{ij} \left[\left(\frac{r_0}{r_{ij}} \right)^{12} - \left(\frac{r_0}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

where N is the number of particles, n is the periodicity, l_i is the bond length, l_0 is the equilibrium bond length, θ_i is the bond angle, θ_0 is the equilibrium bond angle and ω is the torsion angle. The non-bonded interactions are described in terms of the charges of i and j, q_i and q_j, and the separation between the charges, r_{ij}; the rest of the terms are constants.

These terms can be refined further to give a more precise forcefield at extra computational expense. Due to the exponential increase in computational power over recent years, more advanced forcefields are becoming used as standard. However, when considering a forcefield it should be noted that they are empirical and therefore there is no exact form.

2.4) Hybrid quantum mechanical-molecular mechanical approaches

Hybrid Quantum Mechanical-Molecular Mechanical (QM-MM) [25, 26, 29]techniques involve treating the electronically important reacting part of the system using QM where the electrons are represented explicitly, whilst the non-reacting part of the system is represented by using a computationally inexpensive MM forcefield. The total energy of a system, E_{tot}, is calculated in hybrid QM-MM techniques as follows: -

$$E_{tot} = E_{QM} + E_{MM} + E_{QM/MM} \tag{1.18}$$

where E_{QM} is the energy from the parts of the system treated solely by QM, E_{MM} is the energy from the parts of the system treated exclusively by MM and $E_{QM/MM}$ is the energy of the interaction between the QM and MM parts.

The interface between the QM fragment and the MM fragment is hard to represent. There is no problem when considering two separate molecules, one represented by QM and the other by MM. However, problems arise when the system investigated has bonds between the QM and MM fragments. In these systems generally a QM fragment is embedded into a large MM structure, where the whole structure due to computational constraints cannot be studied solely using QM, for example a protein system. In such a case, where the QM and MM regions are in the same molecule, there are bonds bridging the regions that have to be described by both methods. When the whole QM system is calculated, the valences of the frontier QM atoms valences have to be satisfied to get full MO shells. There are various solutions to this problem. A fragment orbital approach for the interfaces between the QM AOs, producing hybrid orbitals collinear to the bond, is one of the approaches used to counteract this. There is one major drawback to this method - the electron density of the whole system needs to be known in advance. Therefore, this method is only useful in special cases. The most widely used method of representing this interaction is the concept of 'link atoms'. These link atoms (essentially hydrogen atoms) are included in the QM fragment using a self-consistent field approach. This is obviously an approximation although the technique is based on intuitive chemical sense. The algorithms used in hybrid QM-MM techniques contain a correction to take into consideration the effect of the link atoms. To reduce the computational expense of some semiempirical QM methods this approach is extrapolated to replace

large side chains with hydrogen atoms. The other effects of the chains in this case are represented using a perturbative approach.

2.5) Methods for sampling phase space

Molecules are dynamic by nature. Even at absolute zero (the minimum temperature possible), they have a zero point energy associated with atomic vibration. For molecules to react or interact they need to orient themselves correctly with respect to one another. Thus the motion of molecules is fundamental to their chemical characteristics. Molecules move due to the forces on their individual atoms. The movement in the molecules can be simulated by the sampling of phase space around the molecule. Two distinct methodologies can be used to sample phase space, Molecular Dynamics (MD) and Monte Carlo (MC) simulations. MD and MC simulation methods are utilised in a diverse range of techniques, which encompass a large area of computational chemistry. A range of texts have been published that detail these approaches. [24-29]The main advantage of these methods over geometry optimisation methods is that they involve extensive sampling of phase space, although this is computationally expensive. MC and MD techniques therefore generally use MM forcefields because of computational constraints. However it is possible to use them in conjunction with QM techniques.

The most common MD technique used is the finite difference method where MD trajectory vectors are generated from a continuous potential. In this technique the time of interaction is divided into fixed time segments and the total force is calculated for each particle in each configuration. The dynamics of reacting systems can be simulated using MC methods. Monte Carlo (MC) is a technique used to produce configurations of a given system using random sampling of the positions of the particles to represent the system. The potential energy is calculated when the particles are in known positions. The MC simulation technique is used to sample throughout $3N$-dimensional space but ignores momentum components unlike the alternative MD simulations method.

Chapter 3
Exploration of Potential Energy Surfaces

A potential energy hypersurface, generally called a Potential Energy Surface (PES), [29, 69]is a description of the energy of a system with the change of one or more degrees of freedom. A potential energy hypersurface shows energy as a function of more than two coordinates, which cannot easily be visualised, whereas a PES shows energy as a function of two coordinates, which can easily be visualised. Due to the problems visualising a multidimensional PES, generally two chosen geometric coordinates are shown, as shown in the example in Figure 3.1. This shows a stepwise mechanism connecting the two minima, or wells, that correspond to the reactant and product states. There are two elementary steps, one from reactants to the intermediate via a first transition state and a second from the intermediate to the products via a second transition state. Most chemical reactions involve more than one step and go through several intermediate structures, local energy minima, and transition states on the reaction path. Each elementary step of a stepwise mechanism has its own transition state and at least one intermediate, whereas a concerted mechanism only involves one elementary step from reactants to the products via a transition state.

The most probable reaction path is the minimum energy path (MEP), the path of least resistance between these two wells. A low gradient 'flat' region on the reaction barrier of the PES corresponds to a saddle region. In the saddle region there is a saddle point at which the species is a transition state, whose nature dictates the rate and direction of the chemical reaction. A transition state (a saddle point on the PES) is a hypothetical thermodynamic state corresponding to a maximum on the minimum energy reaction path between the reaction and product energy wells. In an experimental mechanistic study, the focus is on the properties and structure of a transition state, whereas computationally a transition structure (TS) is derived. The two terms transition structure (TS) and transition state are thus distinct and are non interchangeable. The distinction is that the TS produced using computational methods generally does not incorporate all the thermodynamic properties of the true experimental transition state.

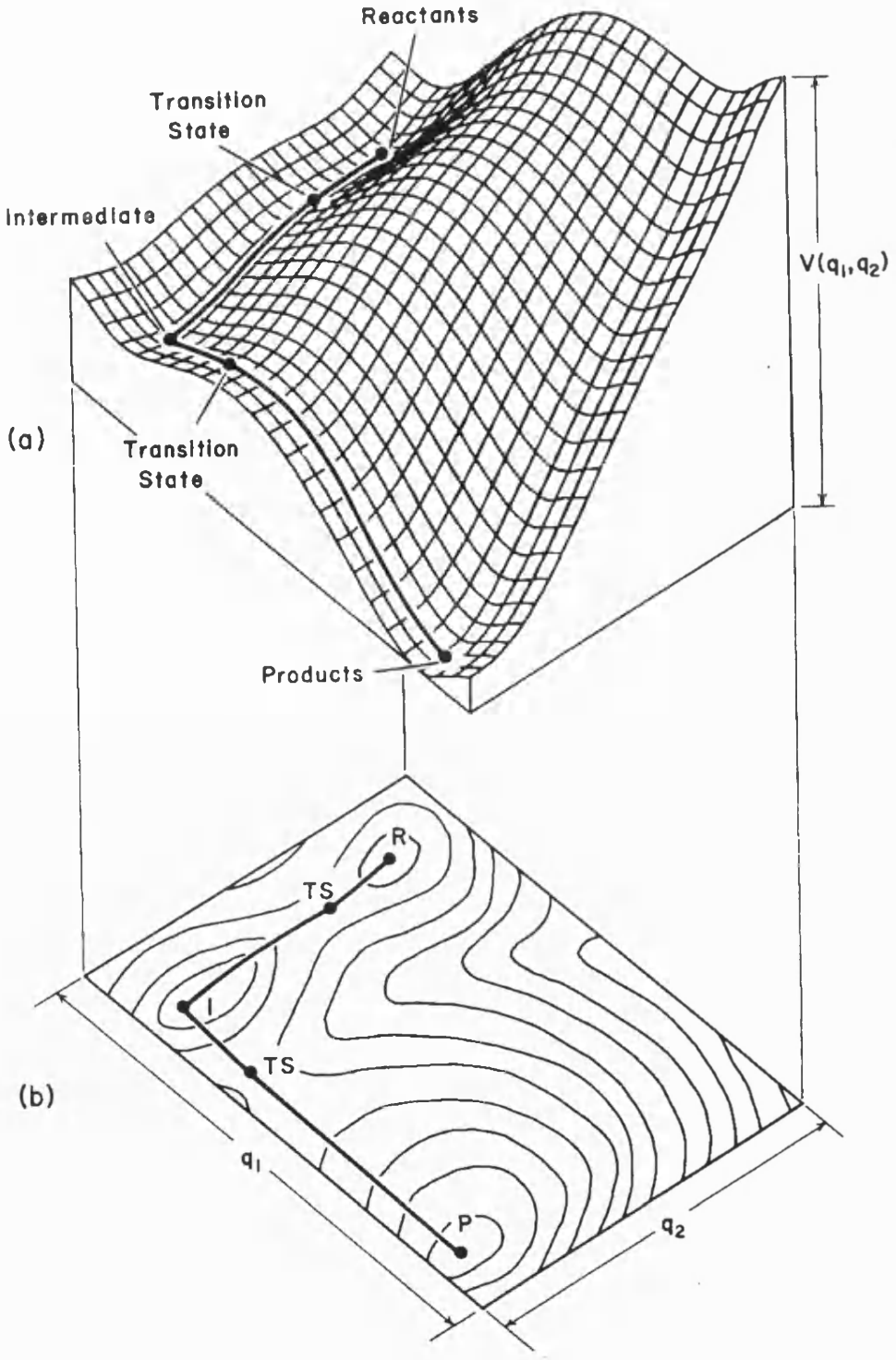


Figure 3.1 – A representation of potential energy surfaces. [70](a) Transect diagram of two-dimensional PES. The heavy dark line of the energy valley runs between the reactant, intermediate, and product points and passes over two TS points as indicated in the Figure. (b) Contour map of a two-dimensional PES. The contours are generated from projecting the transect surface into the q_1 q_2 base plane and represent curves of constant potential energy. The reaction path between the reactant and product points is depicted by a heavy dark line (R $\Rightarrow$ TS $\Rightarrow$ I $\Rightarrow$ TS $\Rightarrow$ P).

Full examination of the entire PES including all the variables, is computationally intractable except for the simplest systems owing to the large number of degrees of freedom. Therefore, most theoretical studies are heavily focused on the small subsection of the PES that is of particular interest, which significantly improves the time taken to solve a calculation. The subsection of interest generally involves a single chemically important reaction coordinate. When studying small systems it is possible to use chemical intuition to predict approximate molecular shapes of the reaction and products, it is also possible to predict the major rearrangements taking place during the reaction. A cross-Section of the PES along a reaction path is called a reaction energy profile; this reaction energy profile contains the chemically significant energy minima and maxima. The reaction energy profile is shown for many of the reactions studied here, as all the chemically important points are present and because a multidimensional energy surface is computationally expensive to compute. It is desirable for the computational method used to accurately describe the PES in solvent as well as in vacuum, as most chemical reactions, including dye fixation, occur in solvent. The inclusion of solvent has been shown to have a substantial effect on the PES; this has been noted in many studies most notably studies on dye fixation. [1]

3.1) Location of stationary points on a Potential Energy Surface

Chemically significant points on a PES generally occur at stationary points where the gradients are zero. The location of these stationary points of a system involves energy optimisation (finding a low energy structure on the surface) and TS search calculations (finding the transition structure, a saddle point on the reaction path), which shall be discussed here.

One issue encountered whilst finding stationary points is the dimensionality of the system and how the atoms are represented. A system with N atoms can be described by 3N mutually orthogonal Cartesian coordinates, or 3N-6 internal coordinates, for which three degrees of translational freedom and three degrees of rotational freedom have been removed. During optimisation the system is often represented in terms of a linear combination of these coordinates. In general the more degrees of freedom a system has the more computationally intensive it is to optimise.

At an energy minimum there is no force acting on the system, which indicates a structurally stable species. Therefore the energy gradients (first derivatives of the energy) are zero and the second derivatives are all positive, with respect to an infinitesimal displacement in each of the coordinates. At a TS the potential energy is at a maximum, with respect to one coordinate and a minimum with respect to all the other coordinates. At this point, the Hessian matrix has a single negative eigenvalue; that is to say the energy gradients with respect to an infinitesimal displacement in each of the normal coordinates are zero and there is one negative second derivative. The direction of the negative curvature indicates the course of the reaction path through the saddle point.

There are various computational algorithms for the exploration of the potential energy of a system, with a large range of publications in the area[29, 70-73]. The methods used to minimise the energy of the structure are undergoing a continuous process of evolution and are well documented in the software packages where they are employed. Therefore, only a brief overview will be given here.

There is no definitive method for optimising the potential energy of a system and so a variety of methods are available within computational chemistry packages. A method that is acceptable for one particular level of theory may not be appropriate for another level. In general the more intensive each individual computational step is the more efficient it is to use a minimisation procedure that uses fewer steps. The methods can be divided into those that use derivatives of the energy and those that do not. Calculation of first and second derivatives of the systems energy can be useful, but depends on the shape of the PES. That is to say if the PES easily fits a parabolic type shape rather than an irregular shape then derivative methods are more appropriate. Minimisation algorithms generally locate the nearest energy minimum to the starting point; however, this minimum may not be a global minimum for the entire system in question. The nearest energy minimum to the starting position may be a local energy minimum for the system. Therefore in order to locate this global minimum, a means of generating a series of starting points is generally required. Some minimisation techniques can make 'up hill' moves in order to break out of the local minimum to find a global minimum, but this can increase the time of the subsequent calculation as it increases the amount of conformational space studied.

Two non-derivative minimisation methods are 'The Simplex method' and 'The Sequential Univariate method'. The Simplex method is based on the manipulation of geometric Figures on the PES to minimise the energy. The Sequential Univariate method systematically perturbs each coordinate and fits a parabola to the points in order to find a lower energy point. These are very rarely used today. In general first derivatives, energy gradients, are calculated and a derivative method is used.

A derivative minimisation method requires the calculation of derivatives with respect to all the coordinates within a system. The majority of QM programs calculate the derivatives analytically, which gives both an exact result and is quicker to calculate than the alternative numerical derivatives. Derivative methods rely on the calculation of the first derivatives, the energy gradient, and second derivatives, the gradient of the energy gradient. There are a number of first order derivative minimisation methods (for example, steepest descent and conjugate gradients) and second order derivative methods (for example, Newton Raphson and Quasi Newton). The sign of the first derivative of the energy indicates the relative position of the minimum and the magnitude indicates the steepness of the local slope. In a minimisation procedure the atoms are moved towards the nearest energy minimum in response to the force, proportional to and in the same direction of the gradient, acting upon them. Second derivatives provide additional information as to where the energy gradient will change direction, that is to say pass through a stationary point, for example a minimum.

The steepest descent and conjugate gradient approaches are first order derivative approaches that work iteratively to converge towards a minimum. The initial point is provided and then each iteration uses the geometry from the previous iteration as a starting point. The steepest descent method moves in a parallel direction to the net force following the steepest route towards the minimum. A line search is then carried out as this may 'overshoot' the minimum. This gets closer to the minimum by interpolating between two points either side of the minimum. This can take a lot of steps therefore a mathematical function (for example a quadratic or a higher polynomial) can be fitted to the surface in order to find the minimum in fewer steps. One drawback of a steepest descent approach is that the rate of convergence slows down towards the minimum and it will actually never reach the minimum. The conjugate gradient method is similar to the steepest descent

method. However, in the steepest descent method both the gradient and successive steps are orthogonal, at right angles to each other, whereas in the conjugate gradient approach the gradients are orthogonal, but the directions are conjugate. Therefore the line searches are performed along a line that is a mixture of the current negative gradient and the direction of the previous search. This makes conjugate gradients a more robust method that is less likely to oscillate, which is a notable problem for a PES that contains 'narrow valleys'.

Second derivative methods use both first and second derivatives of the energy to locate a minimum. The second derivatives of a system are represented as a Hessian matrix or mass weighted force constant matrix. Most second order methods are based on a variant of the Newton-Raphson approach which requires the Hessian matrix to be inverted, as explained later. This is computationally demanding and limits the number of atoms that can be treated by this method to less than about hundred. The Hessian matrix used in Newton-Raphson has to contain only positive eigenvalues; otherwise the method can move uphill, towards a saddle point for example. There are various methods based on the Newton-Raphson approach that do not require the calculation of the entire inverse Hessian matrix. These commonly encountered quasi-Newton methods, for example Broyden-Fletcher-Goldfarb-Shanno (BFGS), Davidson-Fletcher-Powell (DFP) and Murtaugh-Sargent (MS), require first derivatives and construct the inverse Hessian matrix as needed.

A combination of methods is generally used to find a minimum owing to the difficulty of exploring the entire PES. Typically a robust, but less efficient method is used to locate the region which contains the desired minimum, then a more precise method is utilised to accurately and efficiently locate the minimum. QM programs tend to use derivative methods as fewer iterations are required, as the cost of each iteration is much more costly to compute using QM than molecular mechanics techniques.

The Hessian matrix needs to be inverted in order to assist minimisation in second derivative methods. Here is a brief and simplified explanation of why this is the case. An example energy profile for a single coordinate is shown in Figure 3.2. This shows an initial point and the desired energy minimum point for this coordinate with a displacement (Δx) needed in order to move from the initial point to the energy minimum.

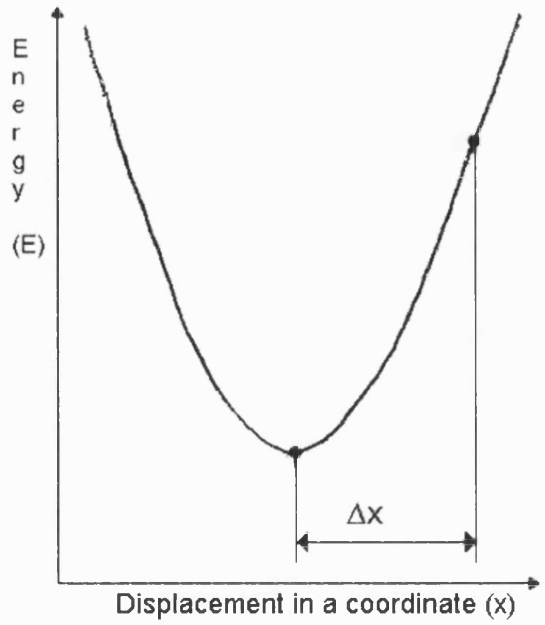


Figure 3.2 – A simplified representation of potential energy surface for a single co-ordinate where the energy varies with a displacement (x). Two points are shown firstly an initial point and secondly the energy minimum point which is a displacement of Δx from the initial point.

The energy (E), first derivative (g) and second derivative for a single coordinate, where H is the force constant for the coordinate and Δx is the displacement from the initial point to the energy minimum are calculated as follows: -

$$E = \frac{1}{2} H(\Delta x)^2 \tag{3.1}$$

$$\frac{\partial E}{\partial x} = g = H(\Delta x) \tag{3.2}$$

$$\frac{\partial^2 E}{\partial x^2} = \frac{\partial g}{\partial x} = H \tag{3.3}$$

In general for a multidimensional system, the gradient, g, is a vector and H is a matrix of the force constants for these vectors. Apart from energy itself these are all vector quantities. The displacement can be calculated by rearranging equation 3.2.

$$\Delta x = \frac{\frac{\partial E}{\partial x}}{H} = \frac{g}{H} = g.H^{-1} \qquad (3.4)$$

The matrix needs to be inverted as it is impossible to divide by a matrix, however the same result is given by multiplying by the inverse of a matrix. Many second derivative energy minimisation procedures require inversion of the Hessian matrix during minimisation, although as it is computationally expensive, it is sometimes calculated as and when it is required.

The location of a saddle point for small systems may be predicted using chemical intuition. In these cases a Grid Search, where the coordinates are systematically varied, can be used to locate the saddle region. Once near the transition structure, in the saddle region, fitting an expression to these points and using calculus can then find the saddle point (TS). One such algorithm used to optimise to a TS is the Berny algorithm which uses a combination of rational function optimisation and linear search steps to find the TS fast and reliably. The initial step is dependent on the coordinate system chosen when the full second derivative matrix is not calculated, therefore it is useful to specify two of the variables so they correspond to the reaction coordinate.

It is possible to use a Coordinate Driving approach to find a Transition State, in this approach one or two significant coordinates are perturbed. This, for example, could involve iteratively lengthening the bond of a leaving group to produce an intermediate structure. The energy of the structure is minimised at each step, whilst the coordinate being stretched is fixed. The highest energy point on this path is an approximation to the saddle point. This produces a series of structures following the dissociation of the leaving group. The Gaussian[46] and MOPAC[74] computational packages contain an algorithm that automatically performs a 'path' calculation, whereas the input needs to be manually adjusted when using AMSOL. [53]

Another systematic method for locating the saddle point is that outlined by Dewar and co-workers; [75]this method involves starting from reactant and product structures and iteratively moving them towards each other. For example, if moving from A to B, where a_i represents the coordinates of A, b_i represents the coordinates of B and i is a particular atom, the reaction coordinate distance, R, between the atoms is: -

$$R = \sqrt{\sum_i (a_i - b_i)^2} \tag{3.5}$$

A is changed into B and B to A by gradually reducing the distance R iteratively. At each point the maximum energy structure is 'fixed' and the structure then undergoes an energy minimisation. The 'fixed' structure is then used as a starting point for the next iteration. This continues until the reaction coordinate distance, R, is small and the two structures will be approximations to the saddle point. When a point is found in the saddle region a variety of methods can be used to locate the saddle point. Some minimisation methods, notably Newton-Raphson, can be used to locate a saddle point; if the initial structure is close to the saddle point, that is to say lying within the quadratic region.

When a reaction is to be modelled, it is much more computationally efficient to minimise the structure at a lower level of theory in order to get closer to the desired position. This avoids intensive calculation of structures around a starting position, which is especially important if this initial point is a long way from the calculated minimum. In this way the structure is refined efficiently, for example the initial structure may be energy minimised using a basic molecular mechanics approach then a semiempirical approach, before it is calculated using density functional theory.

In order to increase the efficiency of a calculation, it is desirable to reduce the degrees of freedom present within a structure. Symmetry can be used to reduce the dimensionality of the system and the number of degrees of freedom present. This reduces the amount of conformational space to search and avoids the calculation of many unnecessary degrees of freedom, for example internal rotations. This can reduce dramatically the amount of time needed to calculate a system, even if it only has C_s or C_2 symmetry. Practically, this entails explicitly relating various degrees of freedom to one another via a specific relationship.

Symmetry can be used in favourable circumstances to transform a TS optimisation into a minimisation, which is computationally a much easier procedure.

3.2) Characterisation of stationary points on a Potential Energy Surface

At a stationary point, the gradients present in the PES with respect to all the coordinates are zero. This condition is the case for energy minima, saddle points and maxima. The nature of a stationary point can be determined from the second derivatives. When using normal coordinates the second derivates of the energy at an energy minimum are all positive, whereas at least one of them is negative at a local maxima. It is therefore necessary to examine the second derivatives (the mass weighted force constant matrix or 'Hessian' matrix) to characterise the structure. This matrix is calculated from the difference in the gradient for a small increment along each coordinate. The Hessian is diagonalised to yield eigenvalues, which are proportional to the square root of the vibrational frequency, and eigenvectors, the normal modes of vibration. For a diatomic molecule, which behaves as a harmonic system, the vibrational frequency, ν, associated with a bond stretch can be derived by:

$$\upsilon = \frac{1}{2\pi}\sqrt{\frac{\kappa}{\mu}} \tag{3.6}$$

where κ is the force constant and μ is the reduced mass of the molecule. The reduced mass for a diatomic is a function of the two atomic masses. In a polyatomic system, the calculation becomes progressively more complex and difficult to solve as the molecular structure becomes more complex. For example, the reduced mass becomes a function of both the molecular geometry and mass of the atoms.

The number of negative eigenvalues (imaginary frequencies) is important for characterising the properties of a particular point. The significance of the number of imaginary frequencies present is as follows.

- Zero - The structure is at an energy minimum. For example, this is the case for the reactants/products structures or a genuine intermediate structure.

- One – The structure is a transition structure for a process (first order saddle point). This could be for the chemical reaction or a conformation change. The magnitude of the transition frequency, $\nu_i^{\ddagger}$, for a chemical reaction has a larger frequency, in the order of hundreds of wavenumbers, than is expected for a conformational transition. If $\nu_i^{\ddagger}$ is very low, it is probably not a chemically significant vibration. It is hard to eliminate very small vibrations especially when investigating a solvated system.

- Greater than one – The structure is a local maximum (second order saddle point). This is the maximum along more than one reaction coordinate and as such is useful only as a precursor to the structure of interest.

If a transition structure is found then, its imaginary frequency needs to be examined to identify the process involved. In order to do this it is useful to visualise its normal mode of vibration. This can be achieved by dividing the vibrational motion into small steps and incrementally moving away from the original, allowing a series of structures to be produced which can be viewed in the same way as a 'movie' for an individual mode.

3.3) Exploration of the reaction coordinate on a Potential Energy Surface

It is not sufficient simply to identify an intermediate or transition state in order to describe a reaction. The whole reaction path from reactant to product needs to be followed. This continuous path starts at the reactants passing through the chemically important point and onto the products. This confirms that the correct reaction coordinate has been examined, and passes from the reactant to the product minima through the chemically important intermediate and saddle points.

This can be achieved by means of an intrinsic reaction coordinate (IRC) calculation, as shown in Figure 3.3.

An IRC calculation starts from the transition structure and traces the reaction path from that point. This algorithm follows a single coordinate along a particular vibrational mode (normal mode of vibration for mass weighted coordinates) that is of interest. Generally, the eigenvector that corresponds to the imaginary frequency of the transition structure is followed along the steepest descent path towards the reactant, product or intermediate structures. An IRC calculation can be helped by adding some additional energy to the system in order for the calculation to overcome any small local energy barrier. This procedure identifies the structures on either side of, for example, a transition structure and identifies any other species along the reaction coordinate.

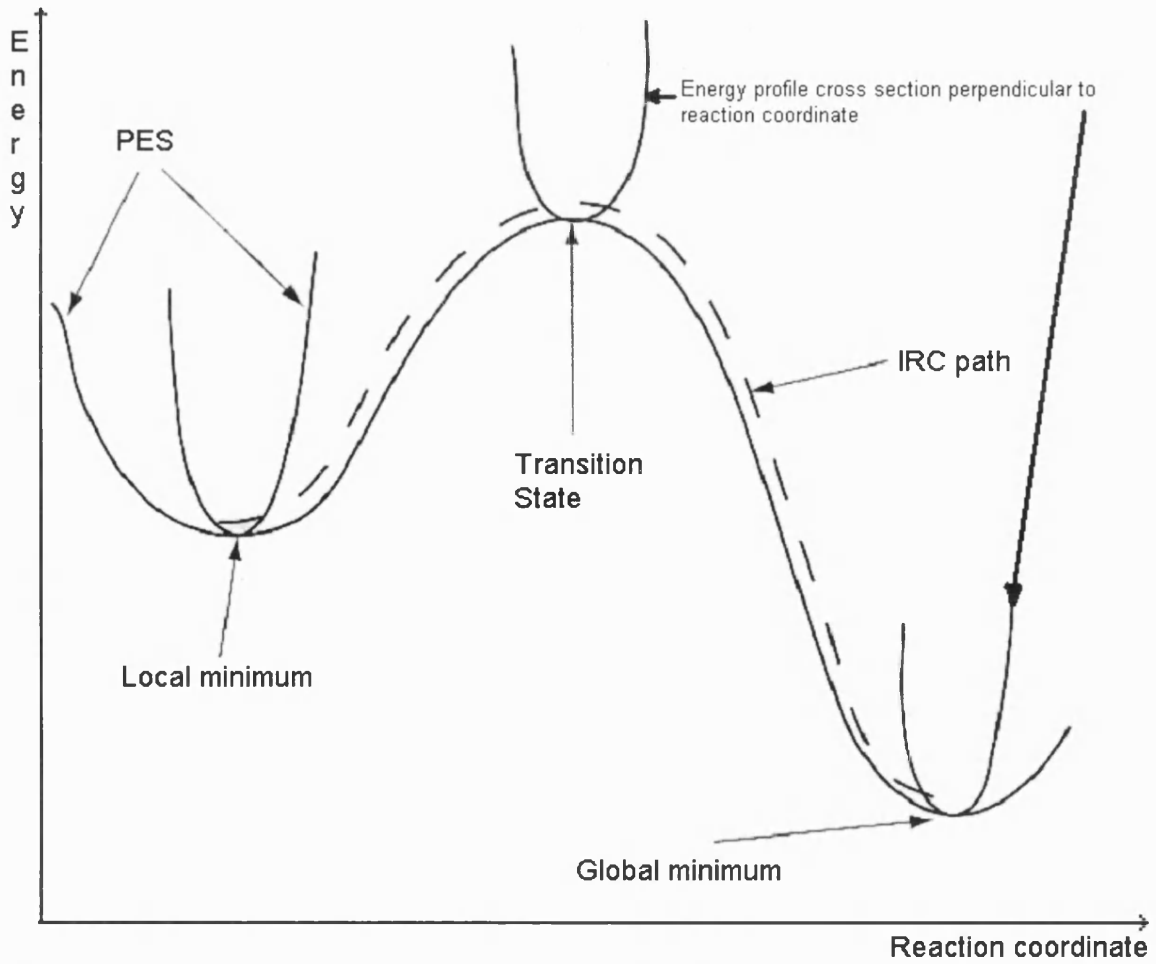


Figure 3.3 – A representation of potential energy surface[76] showing the IRC path between the transition state and both the global minimum and local minimum structure.

It is hard to follow the reaction coordinate when the curvature becomes small at the region corresponding to the transition state or intermediates on the true reaction path. This leads to the inability to locate the transition state along the defined direction and may minimise to the reactant or product valleys. Discontinuities arise when the reaction coordinate does not pass through the saddle point, a discontinuous path does not produce the true maximum with respect to the reaction coordinate. The structures produced during an IRC calculation can be used to visualise a 'movie' of how the reaction proceeds.

Chapter 4

Molecular modelling of dye fixation

What is the mechanism of dye fixation?

4.1) Introduction

The calculations preformed in the studies shown here; were carried out by Dr David Buttar and Mr Philip Hoskin whilst at Bath University. These have not previously been published and provide a basis to the studies undertaken in the remainder of the thesis. Therefore it is reasonable to include these results and analyse and critically evaluate them in this thesis. The evaluation and conclusions drawn have been carried out and inferences drawn by Stuart Gooding, in order to provide a basis for the studies carried out in the further chapters.

The fixation mechanism of reactive dyes, based on the 1,3,5-triazine unit, is to be investigated here using computational techniques. This class of dyes are thought to react by a nucleophilic substitution mechanism, as described in the introductory chapters, with competition between hydrolysis and fixation. The conditions that favour the fixation over the hydrolysis of reactive dyes are of considerable industrial importance; the influences of different leaving groups, nucleophiles, substituents and solvents on the competing processes are to be studied. There are a range of experimental[77-84] and computation studies[85-89] of nucleophilic substitution reactions and experimental studies on the 1,3,5-triazine moiety. [90-94]The nucleophilic aromatic substitution reaction involving the chemically similar nitroarenes, notably trinitrobenzene, has been studied more widely. These have been shown to proceed through the formation of an anionic σ-adduct, Meisenheimer Complex. [7]The reaction of trinitrobenzene with a nucleophile attacking at both the position with the leaving group X (C-1) and ring carbon-3 (C-3) forms two distinct isomeric Meisenheimer complexes here labelled MC-1 and MC-3 respectively, as shown in Figure 4.1. The extent of the formation of the isomeric Meisenheimer complexes is dependant on the nature of the leaving group (X). The formation and decomposition of MC-1 to the substituted product occurs by a nucleophilic aromatic substitution mechanism.

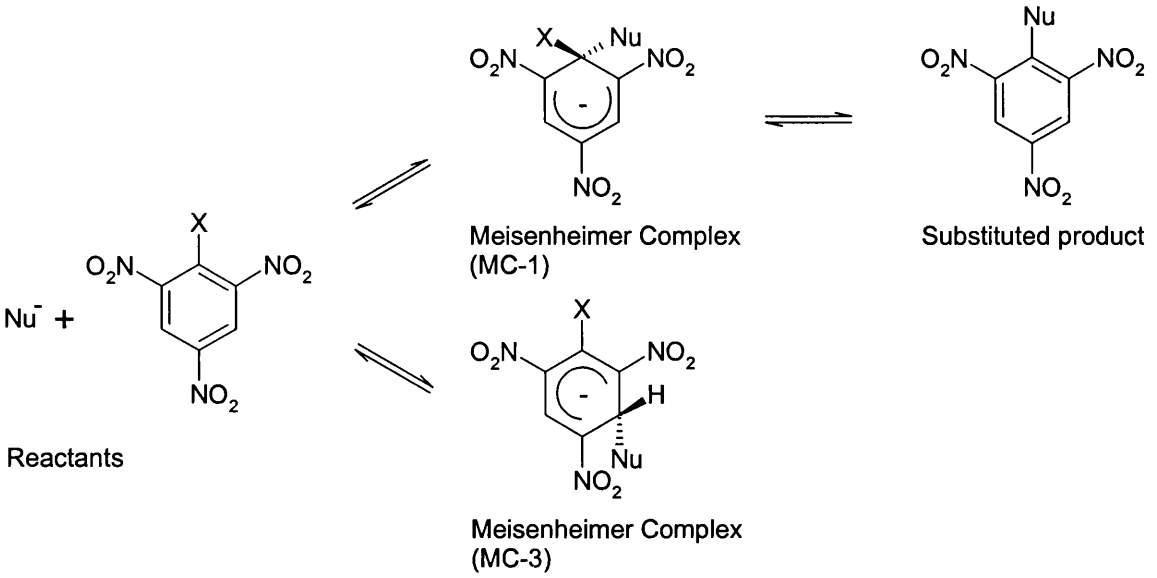


Figure 4.1 – Reaction scheme for trinitrobenzene with a nucleophile (Nu^-) and a leaving group (X).

The concerted or stepwise nature of the nucleophilic substitution reaction leading to the fixation of reactive dyes is of particular interest. A selection of these reactions have been investigated using computational methods to simulate the reactions in both the gaseous and aqueous phases, in order to better understand the fixation process at an atomic level.

4.2) Computational methodology

Semiempirical MO approaches have been used to examine the potential energy surfaces for the 1,3,5-triazine reactive unit. The isolated reactive unit has been studied due to the large computational expense of the inclusion of the remainder of the chromophore. It has been assumed that the chromophore would only slightly perturb the electronic structure of the reactive unit and therefore would not significantly affect the potential energy surfaces produced. The reaction of a dye system binding to a cellulose structure is modelled with a methoxide anion to represent the anionic glucosyl moiety. This approximation is used in many experimental studies and it reduces considerably the computational expense of the equivalent calculations.

The calculations employed the semiempirical AM1 Hamiltonian[40] to simulate the reactions in the gas phase. The addition of a dielectric continuum[52, 54-56, 64-67] was used to simulate the system in the aqueous phase. The solvation models used were the SM2.1 water model[54, 64-67] implemented within AMSOL

5.4.1 program[95] and the COSMO solvation model[52] implemented in the MOPAC 93 program. [74]These dielectric continuum models do not represent the solvent molecules explicitly, but treat solvation purely as a perturbation upon the gas-phase Hamiltonian. In addition to the continuum, the SM2.1 model includes cavity, dispersion and local-field polarisation terms giving the Gibbs free energy of solvation, ΔG_{solv}. The model has been parameterised for both neutral and ionic species in solution, although a lot fewer ionic species than neutral ones were included in the parameterisation set. The root-mean-square errors of the free energies of solvation calculated using the AM1-SM2.1 model is, for example, 3.3 kJ mol^{-1} and 18.4 kJ mol^{-1} for the 148 neutral and 28 ionic species in the parameterisation set, respectively. The calculated gas phase and aqueous phase reactions will be compared and contrasted.

The chemically significant regions of the potential energy surface have been investigated using full geometry optimisation calculations to determine the local and global minima, as described earlier in chapter 3. The Transition Structures (TSs) between these minima have also been located. The gas phase reaction pathways were isolated using an intrinsic reaction coordinate approach, as described in Section 3.3. This is not possible in the aqueous phase; therefore the normal modes of vibration are visualised to characterise the chemically significant structures. For example, for a TS the imaginary vibrational frequency was visualised to see if the movement of the atoms corresponds to movement along the reaction pathway. Additional information about the aqueous TS was found by using a dynamic reaction coordinate approach. [96]

4.3) Results and Discussion

The reactions investigated involve nucleophilic attack on a triazine substrate. The methanolysis (modelling the binding to cellulose) and hydrolysis reaction are of considerable industrial interest in the field of reactive dyes. The substitution reaction at the other carbons is also investigated. The schemes for a stepwise nucleophilic aromatic substitution mechanism for both trinitrobenzene and triazine compounds are shown in Figure 4.2 and 4.3 respectively. The notation shown will be used throughout this chapter, unless stated otherwise. The group on the carbon-1 (C-1) position is denoted X, the group on the carbon-3 (C-3) position is Y and the group on the carbon-5 (C-5) position is Z. The descriptions of the

structures will be shortened to TS for a transition structure and MC for a Meisenheimer Complex. In a reaction with a concerted mechanism the Meisenheimer type complex would be a TS.

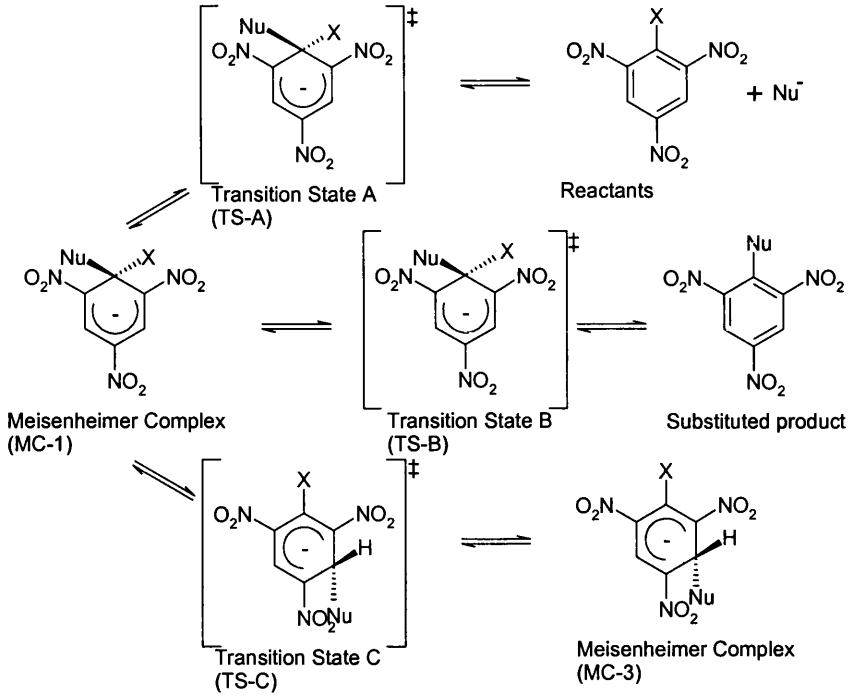


Figure 4.2 – Reaction scheme for the stepwise attack of a nucleophile (Nu^-) on trinitrobenzene with a leaving group (X).

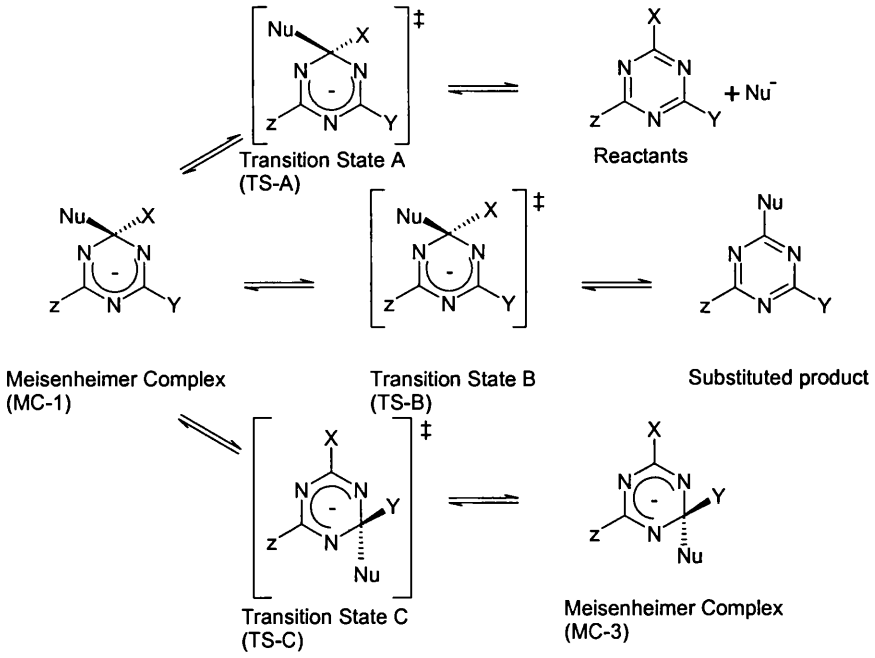


Figure 4.3 – Reaction scheme for the stepwise attack of a nucleophile (Nu^-) on 1-X-3-Y-5-Z-triazine where the preferred leaving group is X.

4.3.1) Hydrolysis and methanolysis of triazines in the gaseous phase

The gas phase enthalpies of formation for the hydrolysis and methanolysis of fluorotriazine and chlorotriazine and hydrolysis of diaminochlorotriazine are shown in Table 4.1. The aqueous enthalpies of formation + Gibbs free energy of solvation ($\Delta H_f + \Delta G_{solv}$) for the same reactions are also shown in Table 4.1. The reaction energy profiles for the gas phase hydrolysis and methanolysis of chlorotriazine and fluorotriazine and hydrolysis of diaminochlorotriazine are shown in Figure 4.4 and Figure 4.5.

The gas phase mechanism for the hydrolysis and methanolysis of triazines and other electron deficient aromatics, [77-84]result in the formation of MC-1, MC-3 and MC-5 complexes. The MC-1 and MC-3 intermediates have been calculated for the reactions in Table 4.1. These MC-1 and MC-3 intermediates can interchange via a TS (TS-C) and then MC-1 can decompose to give products. It has been assumed that the formation of the MC (MC-1) in the gas phase hydrolysis reaction is a barrierless process, as previously reported, [97]therefore (TS-A) has not been calculated. This assumption would be simple to evaluate computationally and could prove not to be valid; for completeness it would be preferable to calculate the TS-A structures. Furthermore the neglect of this calculation does draw into question why the barrier between MC-1 to MC-3 was calculated, which proved to be significant. The methanolysis reactions were shown to have a small barrier, ~12 kJ mol^{-1}, for the formation of MC-1 as found previously. [54] Extrapolating from the magnitude of the methanolysis barrier the hydrolysis may not have a large barrier, but it is possible that there is one.

The diaminochlorotriazine reaction was studied as it is a closer approximation to the structure of a typical reactive dye. This is a large approximation, however, it would be computationally expensive to model a whole reactive dye, which generally encompasses over a hundred atoms. The gas phase hydrolysis of chlorotriazine and diaminochlorotriazine has a similar reaction path. Therefore in the gas phase at the semiempirical (AM1) level of theory chlorotriazine can be used as an inexact but appropriate analogue for a reactive dye.

The existence and structure of TS-C was confirmed by an RHF/6-31G** calculation using GAMESS-UK[98] and it was characterised as a TS by means of a frequency calculation. The geometric structure of AM1 and RHF/6-31G** were in good agreement.

The well depth for the formation of MC-1 in the hydrolysis reactions (A, C and E) was significantly larger, 315, 323 and 280 kJ mol^{-1} than the methanolysis reactions (B and D), 181 and 191 kJ mol^{-1} respectively. This is mainly due to the unduly high enthalpy of formation of the isolated hydroxide anion. The energy calculated for the hydroxide anion used to calculate the energy of the reactants was –59.0 kJ mol^{-1} whereas the experimental value[99] was –161.1 kJ mol^{-1}. The isolated methoxide anion has a smaller error between calculated, -136.8 kJ mol^{-1}, and experimental, [100]-150.6 kJ mol^{-1}.

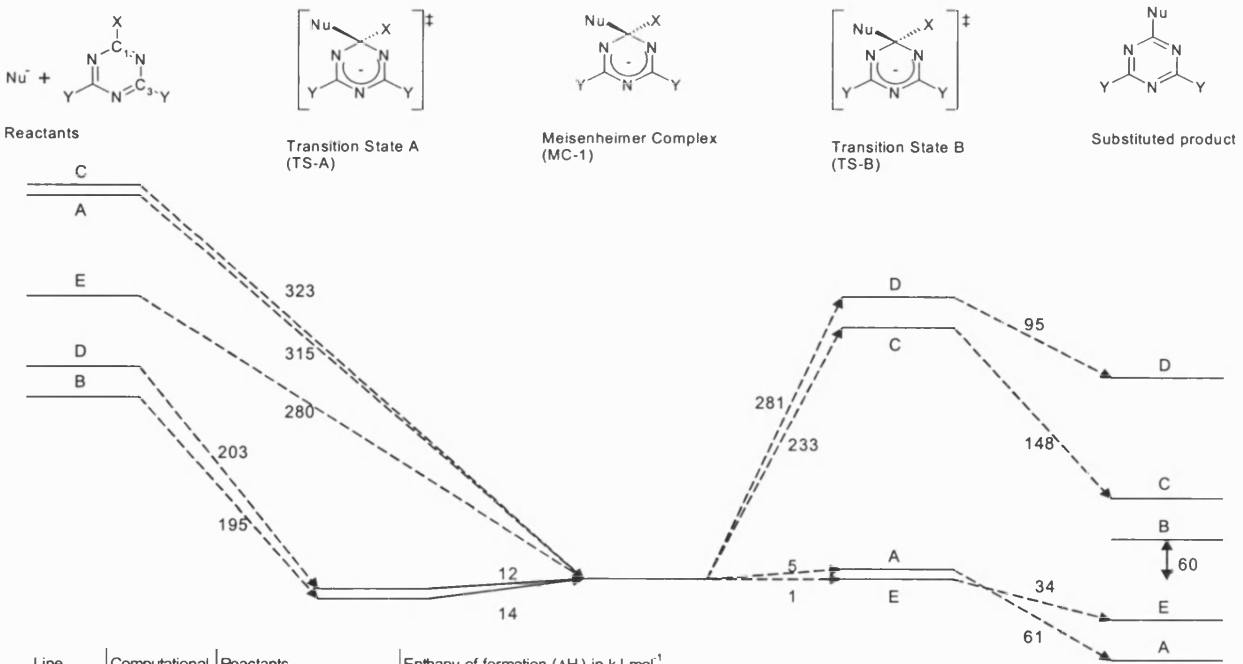
Throughout the gas phase calculations MC-1 was consistently lower energy and therefore more stable than MC-3, as shown in Figure 4.5. There is a large barrier for the interchange between MC-1 and MC-3 via TS-C. This makes interchange less probable and as MC-1 is the thermodynamically preferred structure an excess of products of the C-1 decomposition will be produced compared with the C-3 alternative.

The reactions with a fluoride leaving group have a very high energy barrier for the formation of TS-B for reactions C and D of 233 and 281 kJ mol^{-1} respectively. This is probably due to the energy needed to cleave the strong C-F bond. The reactions with chloride leaving groups readily decompose from MC-1 to products with a barrier of less than 5 kJ mol^{-1} for the formation of TS-B for reactions A and E. TS-B for the other reaction with a chloride leaving group, reaction B, was not found. If reaction B occurs in a similar way to reactions A and E, the small energy difference between MC-1 and TS-B could have contributed to inability to find TS-B. Reactions with the chloride leaving group (A, B and E) occurred with the addition of the nucleophile and the elimination of the leaving group occurring in a single step, where the elimination step is a trivial dissociation to form the product complex. Reactions with a fluoride leaving group (C and D) have significant energy barriers to both steps, the addition of the nucleophile and elimination of the leaving group.

Table 4.1 – The gas phase enthalpy of formation (ΔH_f) and aqueous phase enthalpy of formation with Gibbs free energy of solvation ($\Delta H_f + \Delta G_{solv}$) in kJ mol^{-1} for the hydrolysis and methanolysis of substituted triazines.

Computational method	Reactants				Reactants	Transition State C (TS-C)	Meisenheimer Complex (MC-3)	Transition State A (TS-A)	Meisenheimer Complex (MC-1)	Transition State B (TS-B)	Products
	Substituted triazine			Nucleophile[a]							
	X	Y	Z	$^-$Nu							
MOPAC (AM1)	Cl	H	H	$^-$OH	166.4	105.9	-131.9		-148.4	-143.0	-81.9
MOPAC (AM1)	Cl	H	H	$^-$OMe	64.4	30.4	-96.9	-130.8	-117.0		-57.2
MOPAC (AM1)	F	H	H	$^-$OH	32.3	-37.5	-266.5		-291.1	-57.9	90.0
MOPAC (AM1)	F	H	H	$^-$OMe	-69.8	-111.8	-230.7	-273.0	-261.3	19.4	114.7
MOPAC (AM1)	Cl	NH_2	NH_2	$^-$OH	178.9	125.2	-56.9		-101.6	-100.5	-66.7
AMSOL (AM1/SM2.1)	Cl	H	H	$^-$OH	-343.8	-309.6	-426.9	-307.8	-469.0	-459.9	-471.8
AMSOL (AM1/SM2.1)	Cl	H	H	$^-$OMe	-332.2	-303.9	-380.4	-290.0	-425.7	-417.7	-432.2
AMSOL (AM1/SM2.1)	F	H	H	$^-$OH	-475.2	-441.5	-561.7	-464.4	-604.8	-416.3	-439.3
AMSOL (AM1/SM2.1)	Cl	H	H	$^-$OMe	-421.7			-453.6			
AMSOL (AM1/SM2.1)	Cl	NH_2	NH_2	$^-$OH	-367.2	-308.4	-357.7	-301.1	-445.8	-445.6	-488.3

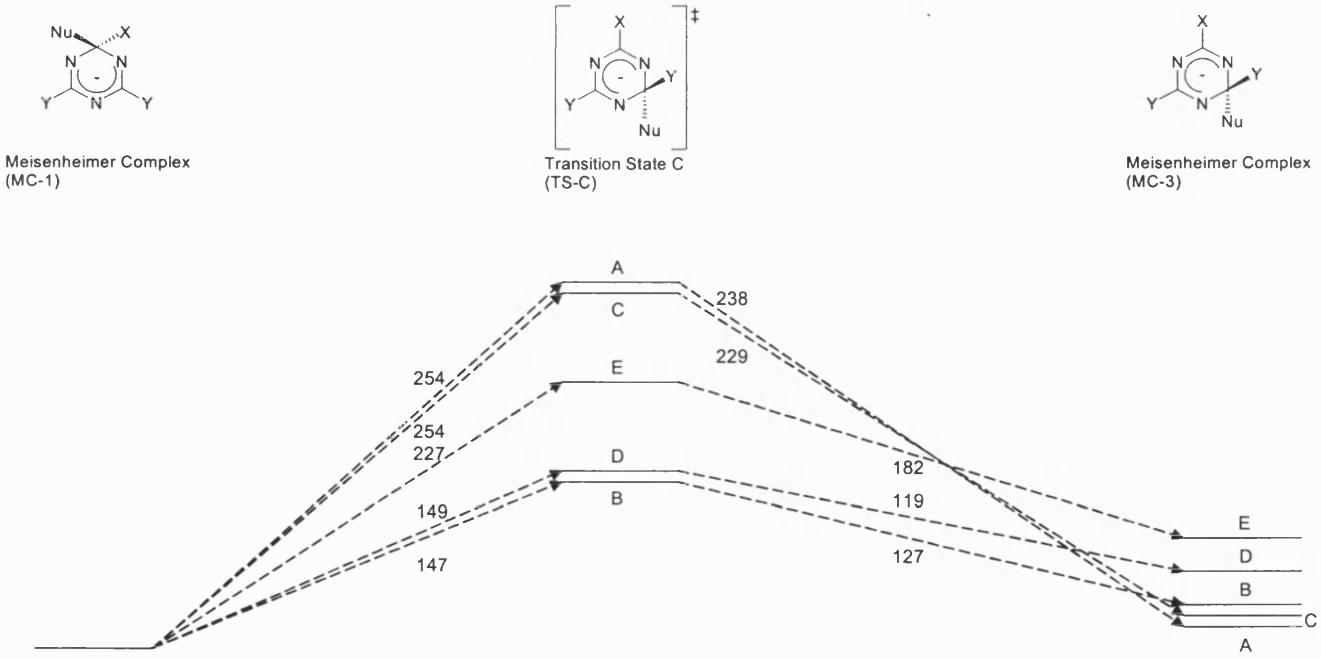
a) The gaseous enthalpy of formation of the hydroxide and methoxide anion was calculated as -59.0 and -161.1 kJ mol^{-1} repectively whereas the experimental values are -136.8 [Lias *et al.*] and -150.6 [West *et al.*] kJ mol^{-1} respectively.



Line	Computational method	Reactants X	Y	Nu[a]	Enthapy of formation (ΔH_f) in kJ mol^{-1} Reactants	TS-A	MC-1	TS-B	Products
A	AM1	Cl	H	$^-$OH	166.4		-148.4	-143.0	-81.9
B	AM1	Cl	H	$^-$OMe	64.4	-130.8	-117.0		-57.2
C	AM1	F	H	$^-$OH	32.3		-291.1	-57.9	90.0
D	AM1	F	H	$^-$OMe	-69.8	-273.0	-261.3	19.4	114.7
E	AM1	Cl	NH_2	$^-$OH	178.9		-101.6	-100.5	-66.7

a) The gaseous enthalpy of formation of the hydroxide and methoxide anion was calculated as -59.0 and -161.1 kJ mol^{-1} repectively whereas the experimental values are -136.8 [Lias *et al.*] and -150.6 [West *et al.*] kJ mol^{-1} respectively.

Figure 4.4 - The gas phase reaction profile relative to the enthalpy of formation of MC-1 for the hydrolysis and methanolysis of chlorotriazine and fluorotriazine and the hydrolysis of diaminochlorotriazine.



Line	Computational method	Reactants X	Y	Nu	Enthapy of formation (ΔH_f) in kcalmol^{-1} MC-1	TS-C	MC-3
A	AM1	Cl	H	$^-$OH	-148.4	105.9	-131.9
B	AM1	Cl	H	$^-$OMe	-117.0	30.4	-96.9
C	AM1	F	H	$^-$OH	-291.1	-37.5	-266.5
D	AM1	F	H	$^-$OMe	-261.3	-111.8	-230.7
E	AM1	Cl	NH_2	$^-$OH	-101.6	125.2	-56.9

Figure 4.5 - The gas phase energy profile relative to the enthalpy of formation of MC-1 for the hypothetical interconversion MC-1 → [TS-C] → MC-3. This is shown relative to MC-1 for comparison with figure 4.4.

4.3.2) Hydrolysis and methanolysis of triazines in the aqueous phase

The same reactions were studied within the presence of solvent using single point MOPAC93[74] AM1/COSMO[52] and fully geometry optimised AMSOL5.4.1[95] AM1/SM2.1. [54, 64-67]These methods were both unsatisfactory for representing the whole potential energy profile, however, AMSOL5.4.1 AM1/SM2.1 was found to be the best of the available methods for investigating the nucleophilic substitution reactions.

The calculated reaction mechanisms could not be described fully using the AMSOL5.4.1 AM1/SM2.1 method as it was not possible to locate and characterise many of the chemically significant points, for example TSs. There is therefore not a complete aqueous energy profile shown for the methanolysis of fluorotriazine. The inability to find these points limits the inferences that can be drawn from the results.

The reaction energy profile for the hydrolysis and methanolysis of chlorotriazine and fluorotriazine and the hydrolysis of diaminochlorotriazine in the aqueous phase are shown in Figure 4.6, and the MC-1 to MC-3 interconversion is shown in Figure 4.7. These are the aqueous equivalent to the gaseous profiles shown in Figure 4.4 and interconversion shown in Figure 4.5. The energies quoted for the aqueous phase calculations are the enthalpies of formation with Gibbs free energy of solvation, as calculated by AMSOL5.4.1 AM1/SM2.1.

The addition of the solvent effects substantially stabilises all the species reducing the energies markedly. The magnitude of the energy difference between the species is also lower. There is however a sizable energy barrier to the formation of TS-A for the hydrolysis reactions and a small barrier was found for the methanolysis reactions.

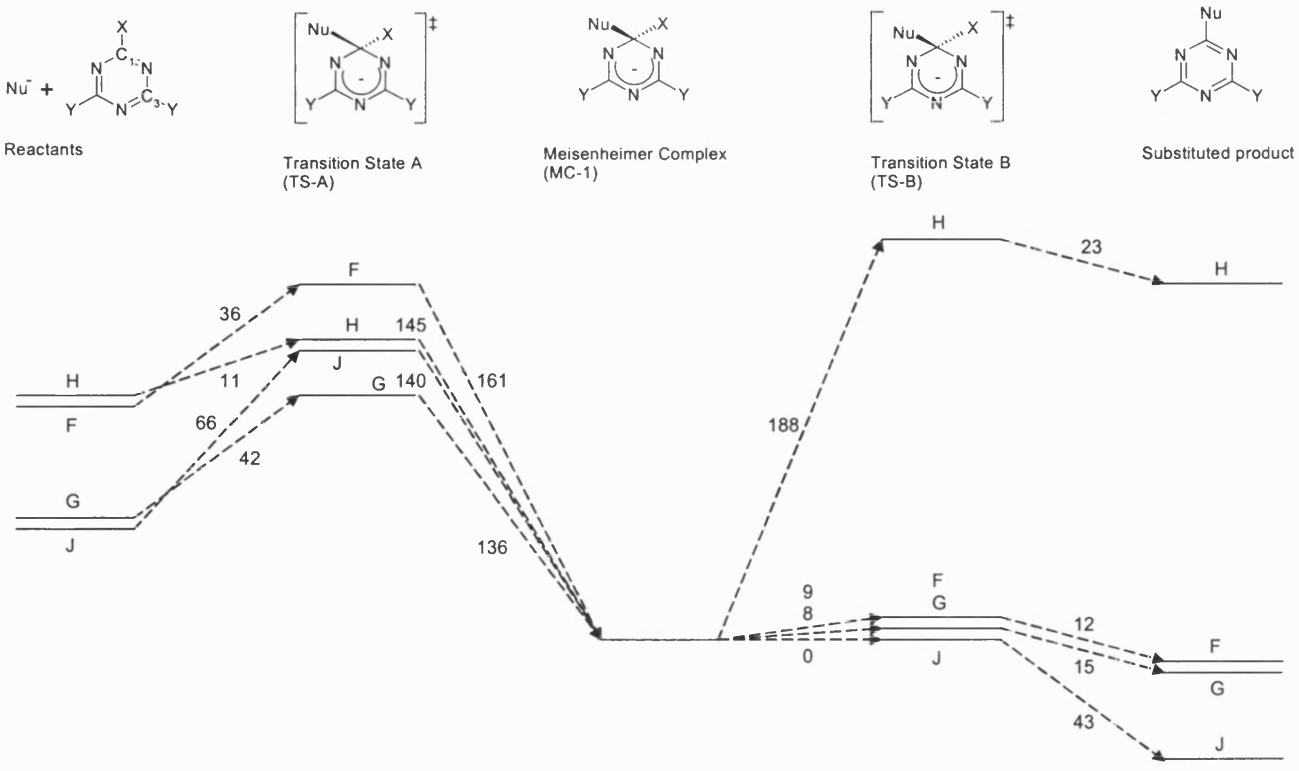
Doubt over the confidence in the data and reliability of the methods are raised due to TS-A being lower in energy than MC-1 for the gaseous methanoylsis of chlorotriazine (B) and methanolysis of fluorotriazine. Therefore either the MC-1 is not sufficiently optimised (highlighting the need for tight convergence criteria) or either the MC-1 or TS-A structures must be incorrect.

The hydrolysis of diaminochlorotriazine (J) was shown to react very much like the hydrolysis of chlorotriazine (F), both having a similar reaction energy profile relative to MC-1. The reactants and products were both comparatively stabilised for reaction J by the amino groups, whereas MC-3 was considerably destabilised. This is probably due to the steric effect of the presence of the amino group in the C-3 position. Due to the similarity of the reaction energy profiles in both the gas and aqueous phase calculations, the use of chlorotriazine instead of diaminochlorotriazine as a model for a reactive dye is a valid approximation.

The well depths for the formation of MC-1 are lower than in the gaseous calculations and are all comparatively similar, ranging from 136 to 161 kJ mol^{-1}. There is no significant difference between the hydrolysis and methanolysis reactions, unlike in the gas phase calculations. This may in part be due to the parameterisation of the solvation model correcting for the difference between the gaseous semiempirical and experimental energies for the hydroxide and methoxide anions.

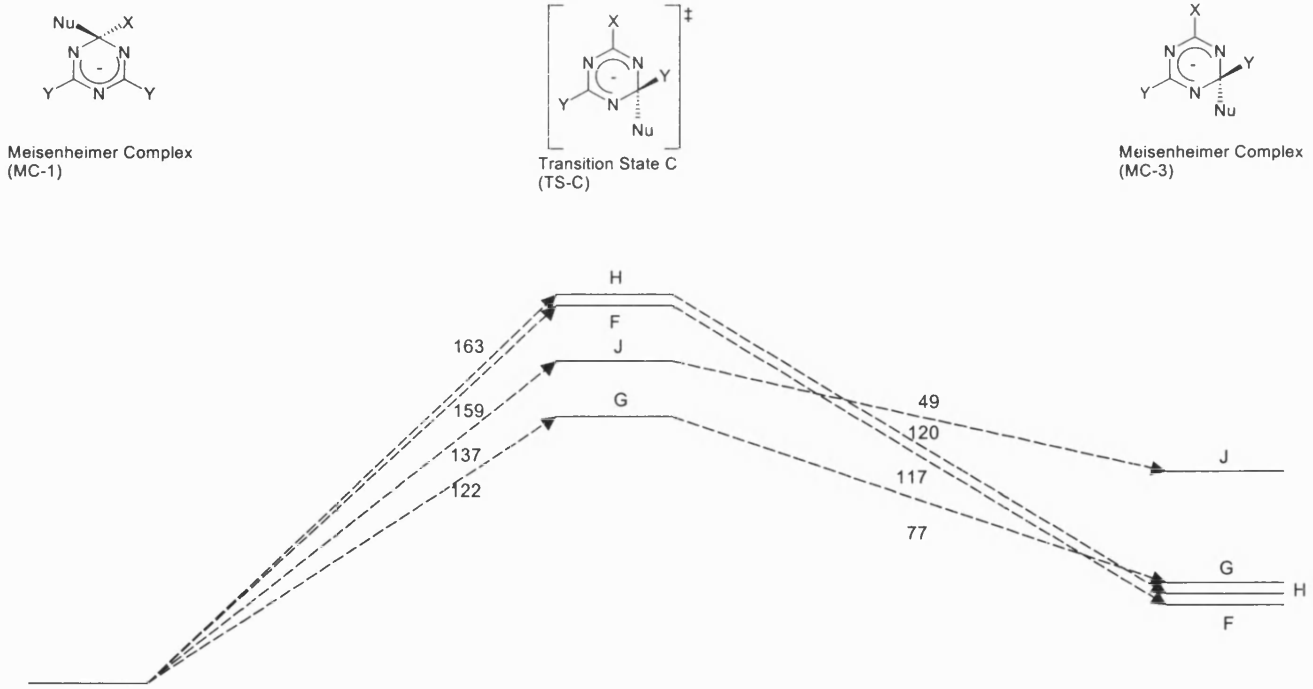
MC-1 is lower in energy than MC-3 for all the aqueous calculations, as in the gaseous phase, and there is a substantial barrier between MC-1 and the formation of TS-C. MC-1 is therefore the preferred lower energy structure and the formation of MC-1 will be the energetically favoured step. Reaction at the C-1 position is therefore more probable than at the C-3 position.

There is a small energy barrier, less than 9.1 kJ mol^{-1}, for the decomposition of MC-1 to TS-B when chloride is the leaving group (F, G and J). The equivalent reaction with fluoride leaving group (H) has a significantly higher energy barrier, 188 kJ mol^{-1}, probably due to the cleavage of the strong C-F bond. Both observations are similar to those made for the gas phase calculations. Therefore both in the gas and aqueous phases the addition of the nucleophile and elimination of the chloride leaving group occurred in a single step, with a trivial dissociation energy barrier; whereas the addition of the nucleophile and elimination of the fluoride leaving group both had significant energy barriers.



Line	Computational method	Reactants X	Y	Nu	Enthapy of formation ($\Delta H = \Delta H_f + \Delta G_{solv}$) in kJ mol^{-1} Reactants	TS-A	MC-1	TS-B	Products
F	AM1 SM2.1	Cl	H	$^-$OH	-343.8	-307.8	-469.0	-459.9	-471.8
G	AM1 SM2.1	Cl	H	$^-$OMe	-332.2	-290.0	-425.7	-417.7	-432.2
H	AM1 SM2.1	F	H	$^-$OH	-475.2	-464.4	-604.8	-416.3	-439.3
I	AM1 SM2.1	F	H	$^-$OMe	-421.7	-453.6			
J	AM1 SM2.1	Cl	NH_2	$^-$OH	-367.2	-301.1	-445.8	-445.6	-488.3

Figure 4.6 - The aqueous phase reaction profile relative to the enthalpy of formation with Gibbs free energy of solvation ($\Delta H = \Delta H_f + \Delta G_{solv}$) of MC-1 for the hydrolysis and methanolysis of chlorotriazine and fluorotriazine and the hydrolysis of diaminochlorotriazine.



Line	Computational method	Reactants X	Y	Nu	Enthapy of formation ($\Delta H = \Delta H_f + \Delta G_{solv}$) in kJ mol^{-1} MC-1	TS-C	MC-3
F	AM1 SM2.1	Cl	H	$^-$OH	-469.0	-309.6	-426.9
G	AM1 SM2.1	Cl	H	$^-$OMe	-425.7	-303.9	-380.4
H	AM1 SM2.1	F	H	$^-$OH	-604.8	-441.5	-561.7
I	AM1 SM2.1	F	H	$^-$OMe			
J	AM1 SM2.1	Cl	NH_2	$^-$OH	-445.8	-308.4	-357.7

Figure 4.7 - The aqueous phase energy profile relative to the enthalpy of formation with Gibbs free energy of solvation ($\Delta H = \Delta H_f + \Delta G_{solv}$) of MC-1 for the hypothetical interconversion from MC-1 → [TS-C] → MC-3. This is shown relative to MC-1 for comparison with figure 4.6.

4.3.3) Hydrolysis of 2-chloro-4-amino-6-N-methylanilio-1,3,5-triazine in the aqueous phase

It has been shown in experimental studies that triazines can react nucleophilically at the C-3 as well as the C-1 position. The hydrolysis at the C-1, C-3 and C-5 positions of 2-chloro-4-amino-6-N-methylanilio-1,3,5-triazine has been studied using the AMSOL5.4.1 AM1/SM2.1 semiempirical continuum approach. The reaction energy profile relative to the energy, enthalpy of formation with the Gibbs free energy of solvation, of the reactant for these reactions are shown in Figure 4.8. All other profiles in this chapter are relative to the energy of MC-1. However as the reactants are the same here, it is sensible to produce the energy profile relative to them.

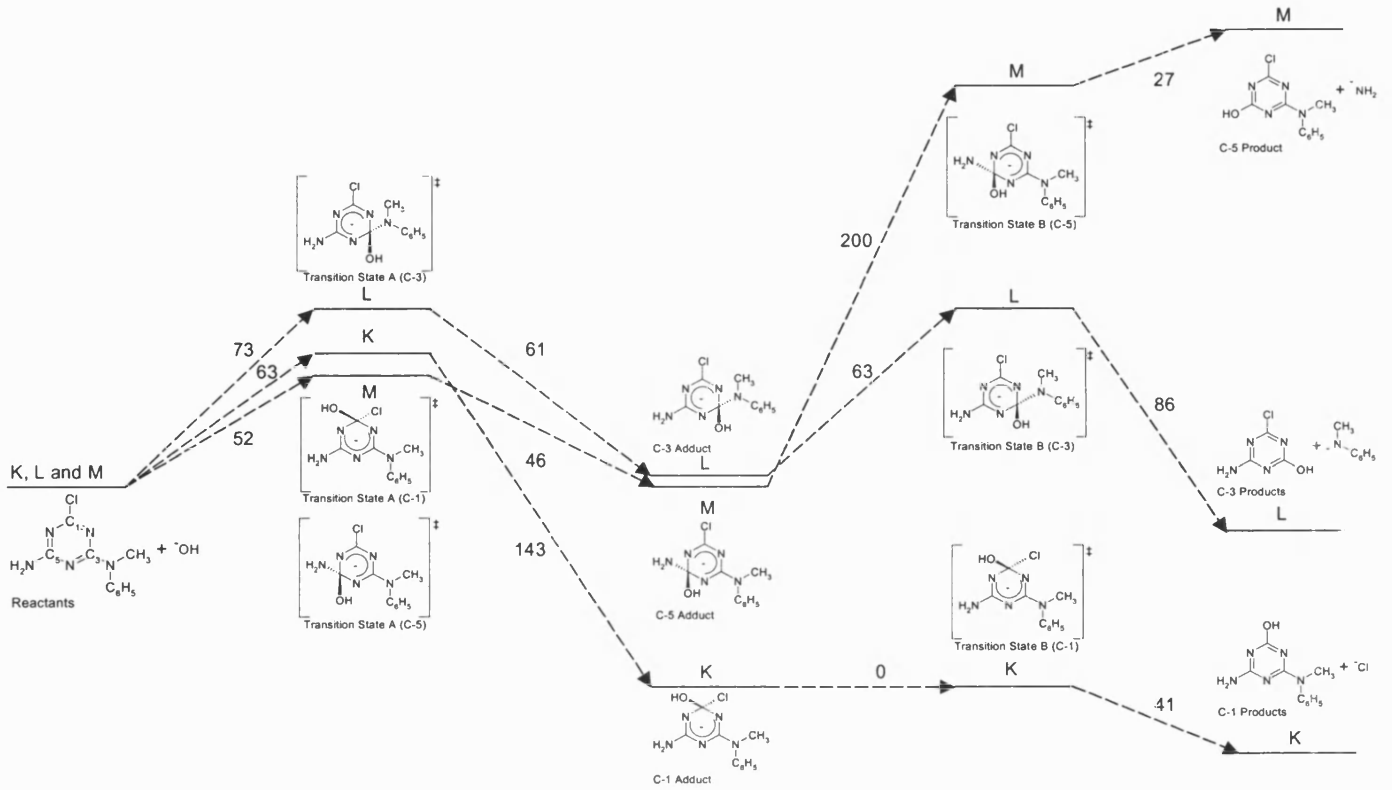
Of the possible active sites the attack at C-5 has the lowest activation energy for the formation of TS-A although there is only 21 kJ mol^{-1} between the reaction at each the possible positions. This is relatively small compared to the magnitude of the energy differences involved in other steps in the reaction.

The C-1 adduct enthalpy of formation, shown in the table of figure 4.8, of -236 kJ mol^{-1}, is significantly lower, over 80 kJ mol^{-1}, than the C-3 and C-5 adducts which have energies of formation of -144 and -150 kJ mol^{-1} respectively. As it is energetically preferred, the formation of the C-1 adduct would be more probable than the other adducts.

There is a substantial difference in the activation energies for the formation of TS-B. The reaction at the C-1 position having a virtually non-existent barrier, whereas the reaction at the C-3 and C-5 positions have barriers of 63 and 200 kJ mol^{-1} respectively. The high activation energies for C-3 and C-5 are due to the poor leaving groups used. The almost non-existent energy barrier for elimination of the leaving group of the C-1 adduct indicates that the attack of the nucleophile and elimination of the leaving group occur in a single step, the nucleophilic attack being rate determining and the elimination being trivial.

The overall reaction is exothermic for attack at the C-1 position, yielding 121 kJ mol^{-1}, slightly exothermic for the C-3 position, yielding 11 kJ mol^{-1}, and highly endothermic for reaction at the C-5 position, requiring 236 kJ mol^{-1}. However, all

the reactions are exothermic if it is assumed that a proton can be taken from the solvent in the final reaction step. This would substantially lower the energy of the products, for example the amino anion would become ammonia and the chloride would become hydrochloric acid. The use of explicit solvation could therefore produce a different energy profile. However, in this case it is likely that all the products would have been stabilised by varying extents and the remainder of the energy profile would have been largely unaffected.



Line	Computational method	Reactants X	Y	Z	Nu	Enthapy of formation ($\Delta H = \Delta H_f + \Delta G_{solv}$) in kJ mol^{-1} Reactants	TS-A	MC-1	TS-B	Products
K	AMSOL (AM1/SM2.1)	Cl	N(Me)(Ph)	NH_2	$^-$OH (C-1)	-156.3	-93.0	-236.0	-236.0	-277.4
L	AMSOL (AM1/SM2.1)	Cl	N(Me)(Ph)	NH_2	$^-$OH (C-3)	-156.3	-83.2	-144.1	-81.3	-167.7
M	AMSOL (AM1/SM2.1)	Cl	N(Me)(Ph)	NH_2	$^-$OH (C-5)	-156.3	-104.4	-150.0	50.0	77.2

Figure 4.8 - The aqueous reaction profile relative to the enthalpy of formation with Gibbs free energy of solvation (ΔH) in kJ mol^{-1} of the reactants for the hydrolysis of 2-chloro-4-amino-6-N-methylanilio-1,3,5-triazine.

4.3.4) Methanolysis, hydrolysis and second hydrolysis of dichlorotriazine and methanolysis of 1-chloro-3-dimethylaminotriazine and 1-chloro-3-dimethylamino-5-nitro-triazine

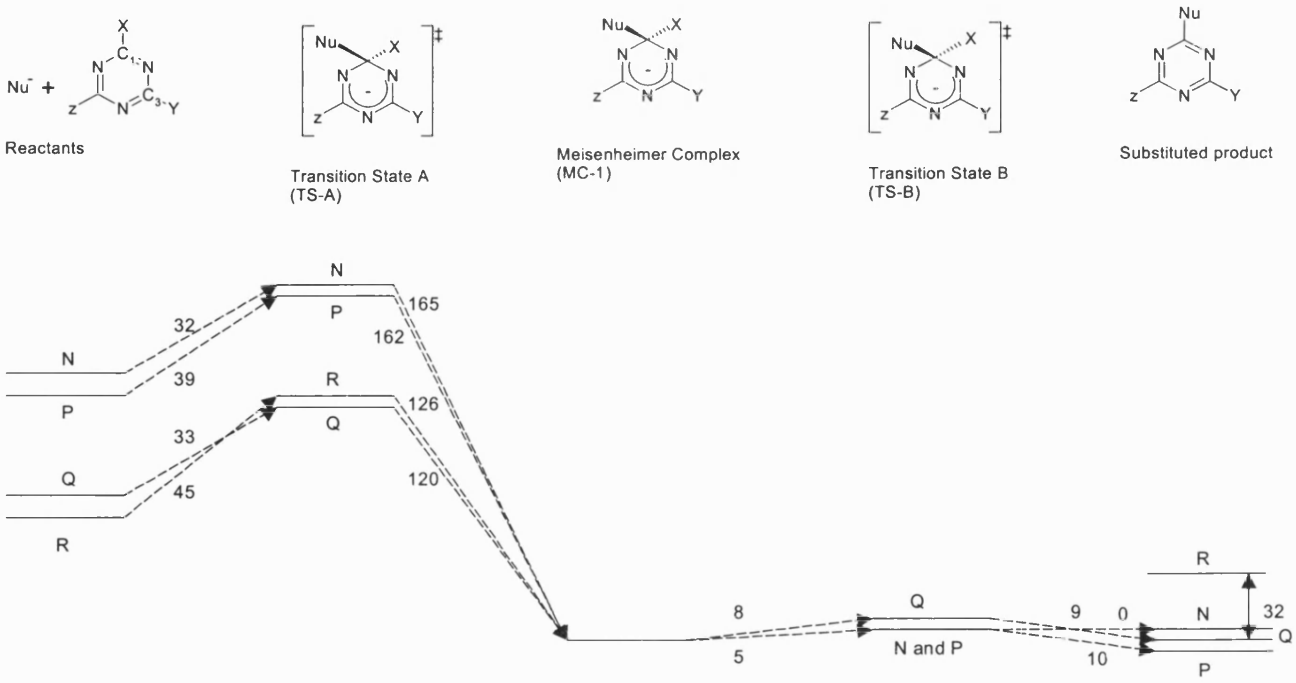
The aqueous energy profiles relative to the energy of MC-1 for the methanolysis, hydrolysis and second hydrolysis of dichlorotriazine and methanolysis of 1-chloro-3-dimethylaminotriazine and 1-chloro-3-dimethylamino-5-nitro-triazine are shown in Figure 4.9. The interconversion between the MC-1 and the only MC-3 found that for the hydrolysis of dichlorotriazine is shown in Figure 4.10. As it was not possible to find all the MCs and TSs for these reactions using the AMSOL5.4 AM1/SM2.1 approach the amount of data to analyse is restricted, for example as MC-1 was not found for reaction O therefore no relative values could be shown on Figure 4.9.

Reactions N and P have approximately the same energy profile, within 10 kJ mol^{-1}; this indicates that the first hydrolysis is similar in nature to the second hydrolysis of dichlorotriazine. This also shows that both chlorine and hydroxide at the C-3 position have a similar effect on the reaction profile. Likewise the methanolysis of 1-chloro-3-dimethyltriazine and 1-chloro-3-dimethyl-5-nitrotriazine, reactions Q and R, have very similar reaction energy profiles. The main difference being that R, with an extra nitro-substituent at the C-5 position, had a comparatively stabilised reactant and destabilised product. Therefore modelling a reactive dye with hydrogen in the C-5 position instead of a nitro substituent is a reasonable approximation. The imaginary frequencies for TS-A for Q and R are 462*i* and 395*i* cm^{-1} respectively clearly showing a significant TS. TS-B for reactions Q and R have however not been possible to characterise as a TS for the reaction coordinate, therefore the energies given are those for approximate structures found.

The reaction all involved a chloride leaving group and like the previous studies have a non-trivial activation energy for the formation of TS-A and a small activation for the formation of TS-B. This makes the reactions occur by addition of a nucleophile and elimination of a leaving group in a single step with a small TS-B barrier and the formation of TS-A being the rate determining step.

Upon examination of the geometries for reactions Q and R, the structures denoted TS-A appear to resemble Meisenheimer complexes with similar bond lengths to the nucleophile and leaving group. The MC-1 and TS-B structures are product-like structures, with the leaving group detached over the triazinyl ring and the nucleophile bonded in the plane of the ring. This may be an artefact of the continuum approach as the leaving group does not interact directly with individual solvent molecules and become subsumed into the bulk solvent.

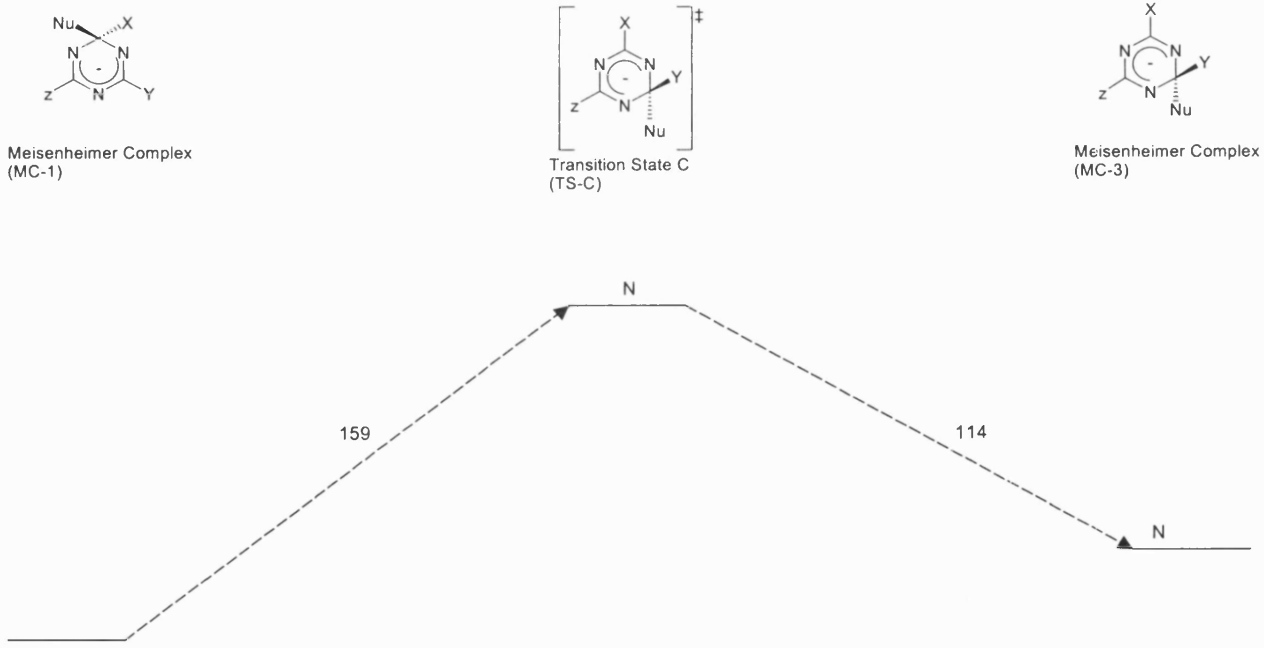
These structures indicate that reactions Q and R are in fact concerted, not stepwise as the energy profile suggests. The Meisenheimer type structure, denoted TS-A, is the TS for this reaction. The spurious TS-B structure may have been formed due to the energy change for breaking an attachment from the π-system of the ring to the leaving group. An artefact of this particular continuum model could cause the formation of the spurious TS-B structure, as mentioned previously. The computational approach to be used in further study therefore needs to be validated.



Line	Computational method	Reactants X	Y	Z	Nu	Enthapy of formation ($\Delta H = \Delta H_f + \Delta G_{solv}$) in kJ mol^{-1} Reactants	TS-A	MC-1	TS-B	Products
N	AM1 SM2.1	Cl	Cl	H	$^-$OH	-346.9	-315.0	-479.5	-474.5	-474.7
O	AM1 SM2.1	Cl	Cl	H	$^-$OMe	-335.2	-299.9		-607.3	-1237.7
P	AM1 SM2.1	Cl	OH	H	$^-$OH	-518.9	-479.7	-641.6	-636.2	-646.6
Q	AM1 SM2.1	Cl	H	NMe_2	$^-$OMe	-258.2	-225.5	-345.6[b]	-337.2	-346.0
R	AM1 SM2.1	Cl	NO_2	NMe_2	$^-$OMe	-217.1	-172.0	-298.3[b]		-266.5

b) TS-A is the Meisenheimer type structure whereas the intermediate appears like the product structure and may be spurious

Figure 4.9 - The aqueous reaction profile relative to the enthalpy of formation with Gibbs free energy of solvation (ΔH) of MC-1, where found, for the methanolysis, hydrolysis and second hydrolysis of dichlorotriazine and methanolysis of 1-chloro-3-dimethylaminotriazine and 1-chloro-3-dimethylamino-5-nitro-triazine.



Line	Computational method	Reactants X	Y	Z	Nu	Enthapy of formation ($\Delta H = \Delta H_f + \Delta G_{solv}$) in kJ mol^{-1} MC-1	TS-C	MC-3
N	AM1 SM2.1	Cl	Cl	H	$^-$OH	-479.5	-320.1	-433.7
O	AM1 SM2.1	Cl	Cl	H	$^-$OMe			
P	AM1 SM2.1	Cl	OH	H	$^-$OH	-641.6		
Q	AM1 SM2.1	Cl	H	NMe_2	$^-$OMe	-345.6[b]		
R	AM1 SM2.1	Cl	NO_2	NMe_2	$^-$OMe	-298.3[b]		

b) TS-A is the Meisenheimer type structure whereas the intermediate appears like the product structure and may be spurious

Figure 4.10 - The gas phase reaction profile relative to the enthalpy of formation with Gibbs free energy of solvation ($\Delta H = \Delta H_f + \Delta G_{solv}$) of MC-1 for the interconversion from MC-1 → [TS-C] → MC-3.

4.3.5) Methanolysis of methoxytriazine and phenolysis of phenoxytriazine using various computational approaches.

The mechanism of an identity reaction, where the nucleophile and leaving group are the same, can be characterised by solely examining the symmetrical species. The symmetrical species is either an equilibrium intermediate structure for a stepwise mechanism or a TS for a concerted mechanism. Identity reactions were the products are the same as the starting materials are of mechanistic interest. These identity reactions are the hardest reactions to study experimentally and the easiest to study computationally. This is due to the symmetrical nature of the species involved that allows a reduction in the amount of degrees of freedom investigated, significantly reducing the computational expense.

The validity of the available computational methods is to be examined by reproducing experimental results. Three different computational approaches AMSOL5.4.1 (AM1/SM2.1), MOPAC93 (AM1) and MOPAC93 (AM1/COSMO) were used, these were the most reliable methods available on two different identity reactions. As all the different approaches have not been applied to all the reactions, the amount and the reliability of the conclusions drawn is limited.

The energy profiles for the phenolysis of phenoxytriazine and the methanolysis of methoxytriazine, the reactions studied are shown in Figure 4.8. It has been shown experimentally that the pyridinolysis of the (4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation[101] and the 4-nitrophenolysis of 4,6-dimethoxy-1,3,5-triazine[91] react with a stepwise (A_N+D_N) and concerted (A_ND_N) mechanism respectively.

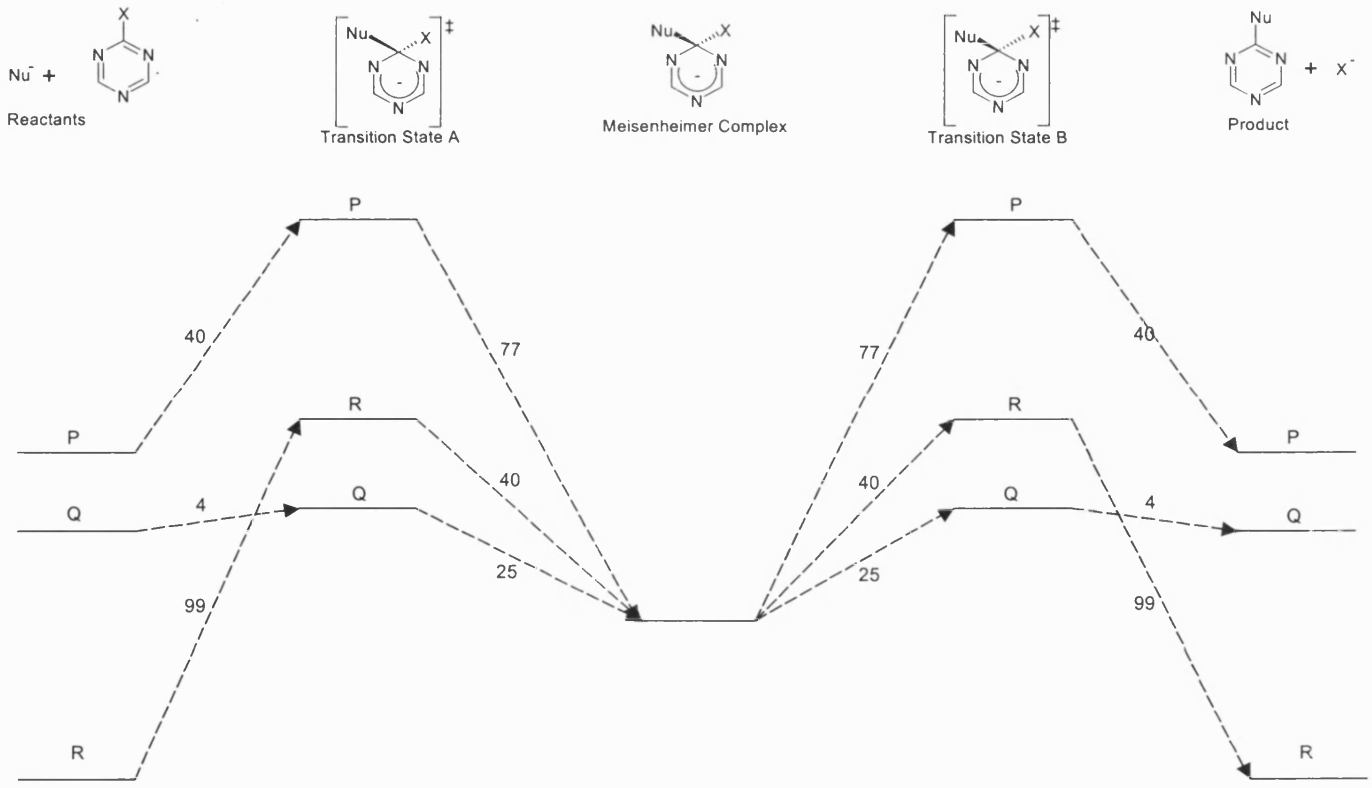
The phenolysis reaction investigated is similar to that of the experimentally studied 4-nitrophenolysis reaction without the additional substituents on both the triazines and the phenolate. Owing to the similarity of these reactants to those used by A. Williams and co-workers[22, 90-92, 101-103] a calculated energy profile corresponding to a concerted mechanism would have been expected. However both the gas phase MOPAC93 (AM1) and aqueous phase MOPAC93 (AM1/COSMO) methods showed the reaction to have a stepwise mechanism. The depth of the energy well for the formation of the intermediate was significant

at 25 and 40 kJ mol^{-1} for the gas phase and aqueous phase calculations respectively.

The methanolysis of methoxytriazine was investigated due to the similar chemical behaviour of the methoxide group to that of the cellulose based fibre, the target of most reactive dyes. This was calculated using the AMSOL4.5.1 (AM1/SM2.1) approach and also produced a stepwise energy profile. This had a deeper energy well for the formation of the MC than for the other identity reactions.

Throughout the three profiles, the aqueous phase calculations have larger activation energy for the formation of TS-A and a larger well depth for the formation of the intermediate.

The TSs and intermediates in all the calculations did not show the spurious geometrical structures observed in the previous Section. The TS geometries indicate that the binding carbon has more sp^3 character than in the reactant geometries. However the triazine ring is also planar allowing for delocalisation and stabilising of the charge across the π-system, presuming that the conjugated ring behaves aromatically. The structure of the intermediate in all the reactions was that of a Meisenheimer-type anionic σ-adduct with identical bond lengths, as expected, either side of the planar triazine ring.



Line	Computational method	Reactants X	Reactants Nu	Enthapy of formation (ΔH_f) in kJ mol^{-1} Reactants	TS-A	MC-1	TS-B	Products
P	AMSOL (AM1/SM2.1)	OMe	OMe	-456.1	-415.9	-492.9	-415.9	-456.1
Q	MOPAC (AM1)	OPh	OPh	100.0	103.8	78.7	103.8	100.0
R	MOPAC (AM1/COSMO)	OPh	OPh	-302.9	-203.8	-243.9	-203.8	-302.9

Figure 4.11 - The reaction profile relative to the energy of MC-1 for the methanolysis of methoxytriazine and phenolysis of phenoxytriazine.

4.4) Overview

The hydrolysis, methanolysis and phenolysis of the substituted triazine reactive unit have been investigated using computational techniques. Dr David Buttar and Mr Philip Hoskin performed the calculations, during their time working at Bath university. The diagrams have been created and analysis and conclusions drawn by Stuart Gooding. This is shown as a basis and starting point for the remainder the work performed in this thesis.

The reactions were modelled both in the gas and aqueous phases. The energy profiles for all the reactions have been calculated as fully as possible, using the available approaches. This allows analysis of the reaction mechanisms within the gas and aqueous phase with different nucleophiles, leaving groups and triazinyl substrates.

The computational approaches used were based on semiempirical AM1 calculations with and without a dielectric continuum to simulate solvation effects. There are lots of other solvation models now available, as described in Section 2.2, for example with explicit solvent molecules. The inclusion of explicit solvent molecules may influence the energy profile given. All of the computational approaches use approximations, due to the constraints on computational expense. Likewise the reactions studied are simplified, for example using the isolated triazine group and using methoxide to simulate a cellulose lattice.

The calculated mechanisms give estimates of the energies and comparative energies of the structures along the reaction coordinate. This allows reaction trends to be found, however, it has been shown that the computational methods used did produce artefacts and should be evaluated. In particular the spurious structures produced in reactions Q and R are misleading where the second transition state is a product-like π-complex. It is therefore necessary to evaluate the methods further reproducing experimentally known mechanisms. When and if a computationally feasible method is found that is able to reproduce experimentally derived mechanism then the challenge of representing dye fixation can be addressed.

4.5) Conclusion

Dr David Buttar and Mr Philip Hoskin performed the calculations presented here. The introduction, diagrams, analysis and conclusions were created and analysed by Stuart Gooding.

The reactions studied showed marked differences between the gaseous and aqueous energy profiles. AM1 was found to describe the energetics of hydroxide and methoxide badly especially the smaller hydroxide anion, this difference was in part corrected in solvent by the parameterisation in the solvation model. All the compounds were stabilised in the aqueous phase notably the ionic species were particularly well stabilised and the energy barriers were smaller in the aqueous phase. As the dye fixation reaction, like most chemical reactions, occurs in the presence of solvent it is important to include these effects in future calculations. These solvated calculations are therefore more chemically significant.

As the same reactions were not studied by multiple methods it is difficult to compare and contrasts the models and draw conclusions about which method is preferable to describe the reactions energetics. However, the ease of use and observed stability of the methods can be examined.

Computationally it was found that the AMSOL5.4.1[95] AM1/SM2.1 solvation model[54, 64-67] was more robust than the alternative MOPAC93[74] AM1/COSMO model[52]. This was the practical observation of optimisation within AMSOL5.4.1 AM1/SM2.1 being preferable to MOPAC AM1/COSMO, this may be due to the most stable treatment of solvation in SM2.1. Therefore more solvated chemical significant points were found using AMSOL than MOPAC, although unfortunately even using this approach it was not possible to find some of the chemically significant points. Due to the difficulty finding optimised structures using MOPAC93 AM1/COSMO the best solvated method used was AMSOL5.4.1 AM1/SM2.1.

Reducing the system to the triazine alone without the presence of groups similar to those on reactive dyes was justified, as there was no significant difference in the reaction energy profiles. This is shown in the similarly of the reaction energy profiles for the reactions with amino and nitro substituents at the C-3 and C-5

positions these were within a few kJ mol^{-1} of the profiles with hydrogen at the C-3 and C-5 positions. This considerably reduced the computational expense of involved.

In all the non-identity reactions the rate determining step was the attack of the nucleophile on the triazine reactive unit leading to the formation of the Meisenheimer intermediate. The decomposition of MC-1 to form products had a comparatively small barrier. This decomposition barrier was higher when a less favourable leaving group was used. For example the chloride leaving group had virtually no barrier, a few kJ mol^{-1}; however the fluoride leaving group had a comparatively high barrier, ~160 kJ mol^{-1}, due to the energy needed to cleave the strong C-F bond. The reactions with a chloride leaving group involved the attack of the nucleophile and the elimination of the leaving group occurring in a single step with a comparatively trivial elimination step and those with a fluoride leaving group had a significant elimination step. Therefore the reactions with a fluoride leaving group were much more stepwise.

All the reactions, except for the methanolysis of 1-chloro-3-dimethylaminotriazine and 1-chloro-3-dimethylamino-5-nitrotriazine, reacted using a stepwise addition-elimination (A_N + D_N) mechanism. The exceptions show product like structures after the formation of a Meisenheimer type TS. These reactions appear to occur using a concerted (A_ND_N) mechanism. The spurious product-like structures produced could be due to artefacts caused by the particular solvation approach used, implying an error in the computational method. Alternatively this is a genuinely new mechanism.

The studies here have shown that the inclusion of solvation effects is necessary when modelling a dye-based system. The unexpected structures created for two of the reactions have cast doubt on the reliability of the methods used. Therefore the available methods need to be evaluated against experimental known mechanistic data for triazine based reactions.

Chapter 5
Modelling nucleophilic aromatic substitution mechanisms for derivatives of triazine and trinitrobenzene

Can semiempirical continuum methods reproduce experimentally derived mechanisms?

5.1) Introduction

The chemical reactions investigated involve triazine rings and the trinitrobenzene analogues, as these are important in the fixation of dyes. The computational techniques used need to be able to reproduce the experimentally derived mechanisms. Here, three reactions are examined using both the AM1 and PM3 semiempirical Hamiltonians, with two different continuum solvation models.

The reaction mechanisms for two triazine systems with a range of aromatic nucleophiles have been studied experimentally by A. Williams and co-workers at the University of Kent, in collaboration with Zeneca Specialities, [22, 92, 102-105]by means of extended Brønsted correlations as described below in Section 5.1.1. As shown in Figure 5.1, the RHF/6-31G* calculated electronic charge distributions for the aromatic rings of triazine and trinitrobenzene are similar, since the endocyclic nitrogen atoms of the former and the exocyclic nitro substituents of the latter both have electron-withdrawing characteristics. Nucleophilic aromatic substitution of trinitrophenyl compounds has been studied much more extensively than for triazinyl compounds, and recently an analogous extended Brønsted correlation was published by Crampton and co-workers at the University of Durham. [106]

Figure 5.1 – Electrostatic potential (red=negative, blue=positive) mapped onto the electron density surface of trinitrobenzene and triazine calculated using HF/6-31G* in the gas phase.

The ideal computational method should be efficient and able to reproduce experimental results accurately. In order to evaluate the semiempirical continuum methods, the following three identity reactions have been investigated, corresponding to the reactions previously studied experimentally.

Reaction 1. 3-nitrophenolysis of 3'-nitrophenyl-2,4,6-trinitrophenyl ether (Figure 5.2). Experimentally this reaction has been inferred to proceed by a stepwise (A_N + D_N) mechanism involves a Meisenheimer intermediate and two transition states. [106]

Reaction 2. Pyridinolysis of 4,6-diphenoxy-1,3,5-triazin-2-yl pyridinium cation (Figure 5.3). Experimentally, this reaction has also been inferred to proceed by a stepwise (A_N + D_N) mechanism and involves a Meisenheimer intermediate and two transition states. [103]

Reaction 3. 4-nitrophenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Figure 5.4). This reaction has been inferred experimentally to proceed by means of a concerted (A_ND_N) mechanism, involving only a single transition state. [92]

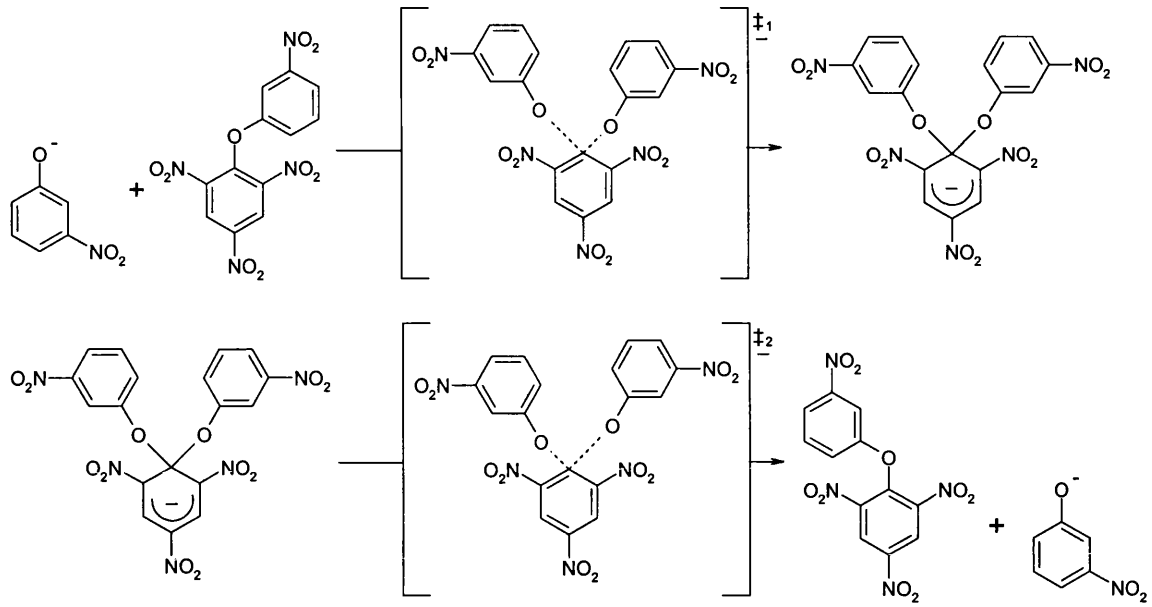


Figure 5.2 – The stepwise (A_N + D_N) mechanism for the 3-nitrophenolysis of 3'-nitrophenyl-2,4,6-trinitrophenyl ether.

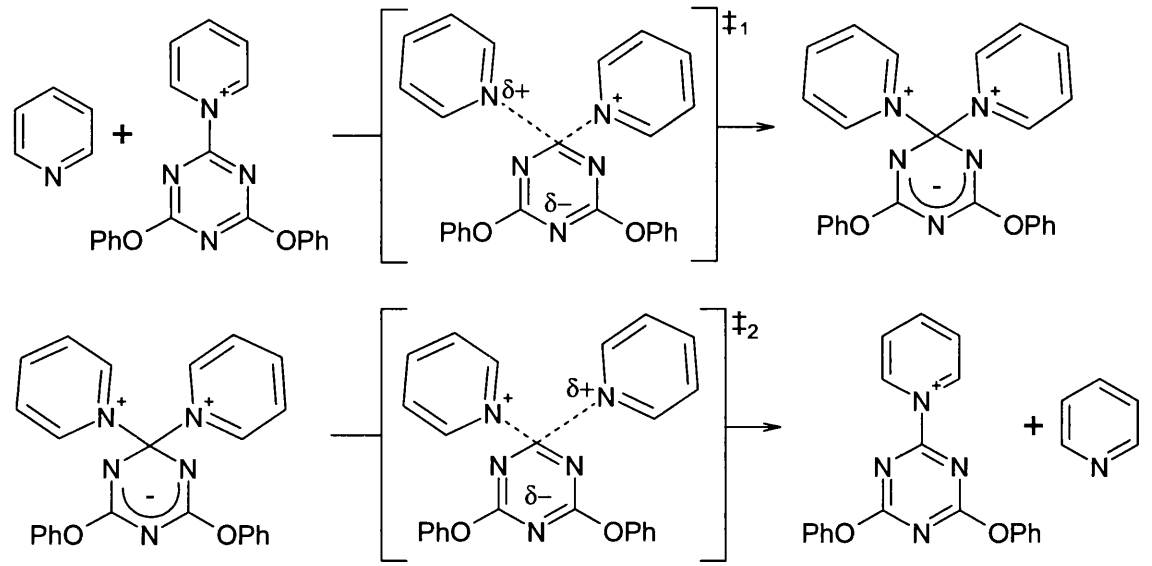


Figure 5.3 – The stepwise (A_N + D_N) mechanism for the pyridinolysis of the 4,6-diphenoxy-1,3,5-triazin-2-yl pyridinium cation.

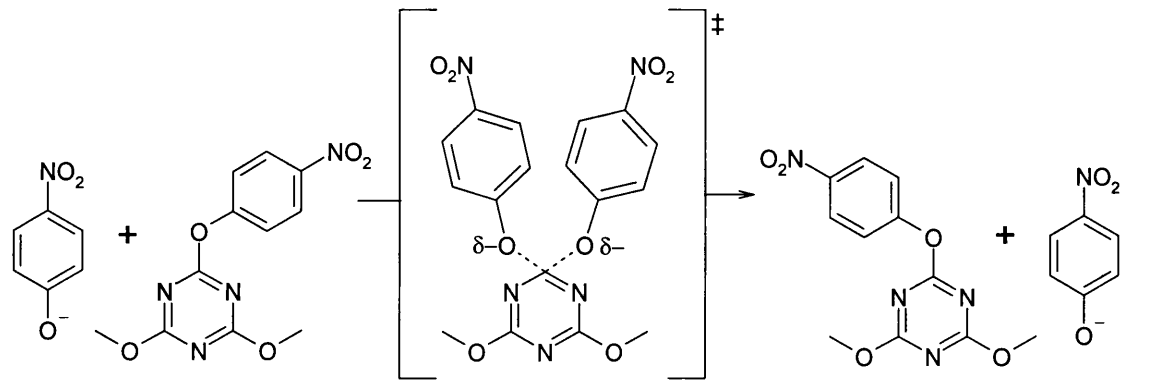


Figure 5.4 – The concerted (A_ND_N) mechanism for the 4-nitrophenolysis of 4,6-dimethoxy-1,3,5-triazine

5.1.1) Experimental Methods

Linear free energy relationships are a popular experimental technique in studies of reaction mechanisms. The slopes of these empirical correlations are often considered to give information about the transition state. One of the most important and regularly studied relationships is the extended Brønsted correlation, which relates the logarithm of the observed reaction rate constant k_{obs} to the equilibrium basicity of a nucleophile as measured by the pK_a of its conjugate acid, as in equation 5.1.

$$\log_{10} k_{obs} = \beta_{nuc}\, pK_a + C \qquad (5.1)$$

The coefficient β_{nuc} is the Brønsted exponent for a family of reactions involving a range of nucleophiles, whose value is often considered to be related to the degree of bond formation at the TS.

A set of reaction energy profiles is shown in Figure 5.5, the first step is sensitive to the nature of the nucleophile and a second step is insensitive. By varying the substituent on the nucleophile to increase its nucleophilicity/basicity, and thus raise the pK_a of its conjugate acid, the barrier for the first step is diminished. As long as the energy for TS_1 at the top of the first barrier remains higher than that of TS_2 at the top of the barrier for the second step, then k_{obs} is determined by the properties of TS_1 for the rate-limiting first step. A plot of $\log_{10} k_{obs}$ against pK_a shows a straight line with a positive slope (Figure 5.6). When the energy for TS_1 falls below that of TS_2 then the rate-limiting step changes; now k_{obs} is determined by the properties of TS_2. If this second step is completely independent of the nucleophile (as implied by Figure 5.5) then the extended Brønsted plot for higher values of pK_a is a straight line of zero gradient (Figure 5.6). The break point in the plot between the two straight-line Sections of different slope corresponds to a reaction in which the energies of TS_1 and TS_2 are equal. This situation would arise when the nucleophile and the leaving group were the same; in other words, for an identity reaction. However, if a range of nucleophiles is considered whose conjugate acids have pK_a values both higher and lower than that of the leaving group, but the extended Brønsted plot shows only a single straight line without a break-point (Figure 5.7), then this is usually interpreted as evidence for a concerted mechanism involving only one TS.

An identity reaction is hard to study experimentally, because the reactants and products are the same. However, due to this symmetry it is easy to study computationally. Important mechanistic information can be found simply by characterising the symmetrical species involved in an identity reaction. An energy-minimum species implies a stepwise mechanism with an intermediate, but a TS implies a concerted mechanism.

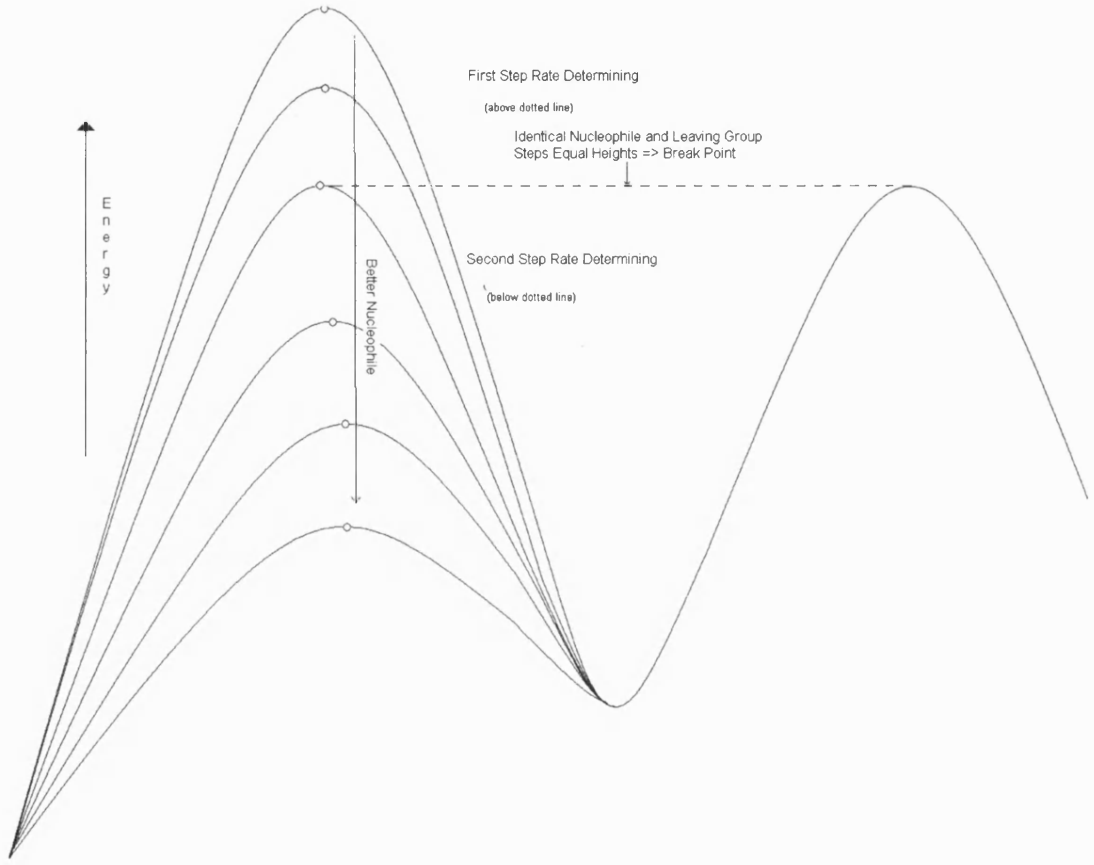


Figure 5.5 – Schematic representation of how reaction energy profiles change depending on the nucleophile used. When a better nucleophile is used the rate determining step changes from being the first step to being the second step via the identity reaction where both the steps and nucleophile/leaving group are the same.

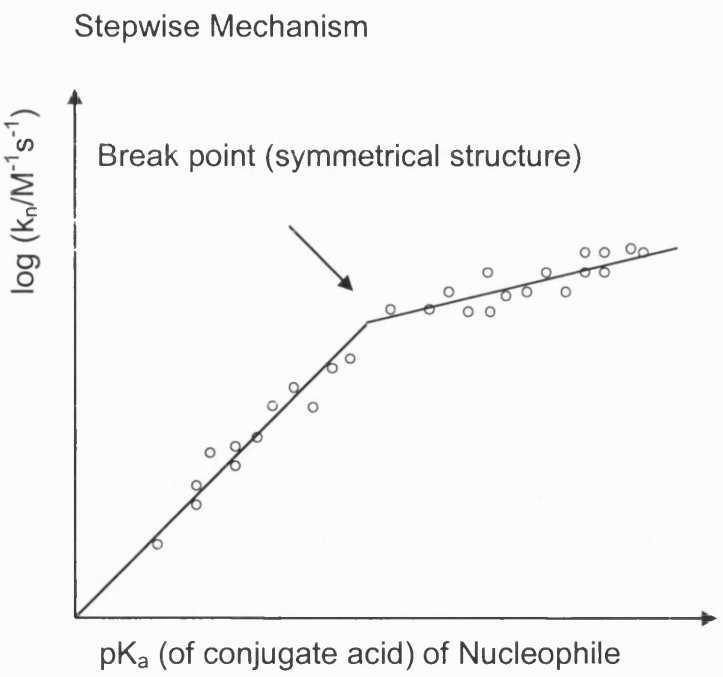


Figure 5.6 – Brønsted of the log of the rate constant (k_{obs}) against the pK_a of the conjugate acid of the nucleophile for a stepwise mechanism.

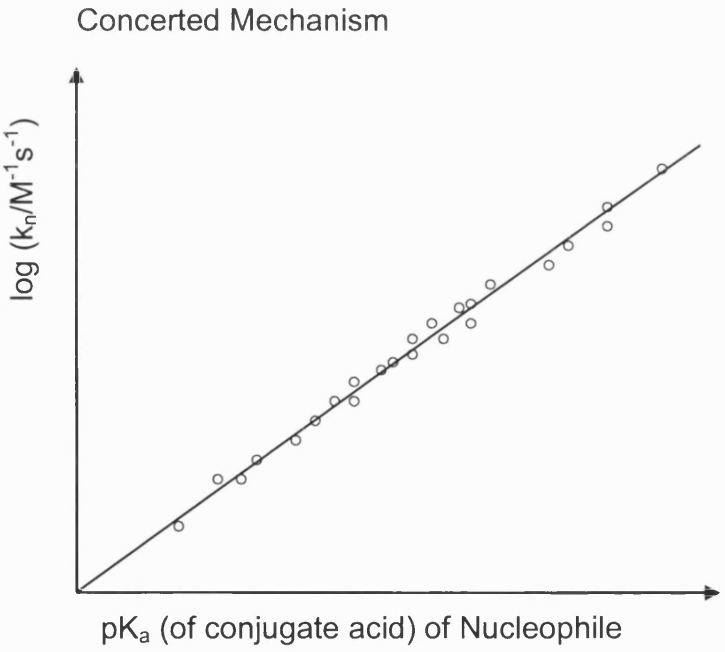


Figure 5.7 – Brønsted of the log of the rate constant (k_{obs}) against the pK_a of the conjugate acid of the nucleophile for a concerted mechanism.

5.1.2) Computational Methods

The methods described in Chapter 3 have been employed to study identity reactions, in which the nucleophile and leaving group are the same. The geometry of a symmetrical species, in which the bonds to the nucleophile and leaving group are constrained to be equal, is optimised by energy minimisation. A frequency calculation is performed to characterise this species as either an intermediate or a transition structure. If it is an intermediate, then one bond is systematically lengthened until an energy maximum is found, at which an "OPT=TS" and frequency calculation are then performed. The reactant and product structures are located by unconstrained energy minimisation

In order to provide continuity with previous modelling studies, the same molecular orbital methods have been employed, namely AM1[40] and PM3, [45]both in the gas phase and with a dielectric continuum used to represent aqueous solvation (ε = 78.4). The aqueous enthalpies of formation were calculated using the COSMO[52] continuum within the MOPAC93[74] program and the aqueous enthalpy of formation with Gibbs free energy of solvation were calculated using the SM5.4A[107] and SM5.4P[107] solvation models within AMSOL6.1. [53](SM5.4A and SM5.4P implications are different parameterisations of SM5.4 optimised for AM1 and PM3 respectively.). The solvation model is an additional source of error in the calculations. The root mean square errors for the parameterisation set was 2.1/1.7 (SM5.4A/SM5.4P) kJ mol^{-1} for the 215 neutral species and 18.0/18.4 (SM5.4A/SM5.4P) kJ mol^{-1} for the 34 ionic species. Similar errors are present in COSMO. There is a significant increase, nearly an order of magnitude, in the computer time needed for an aqueous system compared to a gas phase simulation.

The reactants and intermediate species were each described using a Z-matrix, created either manually or by use of the Molden program, [108]to provide an initial starting point for the geometry optimisation. The following keywords were used to extend the convergence criteria substantially over the default: GNORM=0.0 (MOPAC), GCOMP=0.0 (AMSOL), KICK=3 (AMSOL), DDMIN=0.0, SCFCRT=1E-10 and HESS=3. The eigenvector following optimiser was used where possible, in the COSMO calculations as advised. [52]

5.2) Results and discussion

All energies stated here are in kilojoules per mole (1 kJ mol^{-1} = 0.2390 kcal mol^{-1}). The GNORM, gradient norm, of an optimised structure is a measure of the precision of the optimisation; this is a function of the root mean square deviation of the gradient. The gradient norm has units of kilojoules per reduced unit, where one reduced unit is an Å or a radian depending which internal coordinate is being investigated. This is larger, by a factor of 4.184, than the normally quoted kilocalories per reduced unit. The number of imaginary frequencies, NIMAG, indicates the nature of the structure found; all the frequencies stated are quoted in wavenumbers (cm^{-1}).

5.2.1) The 3-nitrophenolysis of 3'-nitrophenyl-2,4,6-trinitrophenyl ether (Reaction 1: Experimentally stepwise)

Alternative nucleophilic aromatic substitution mechanisms for this reaction are shown in Figure 5.8. Reaction 1 was investigated using the AM1/COSMO and AM1/SM5.4A computational approaches only. Three reaction profiles have been calculated because the "symmetrical" intermediate has three conformers, due to the flexibility of the phenoxy link to the substrate; these are shown in Figures 5.9, 5.10 and 5.11. Diagrams and energies of the three 'symmetrical' intermediate are shown in Figure 5.12. The structures A, B and C have C_1, C_2 and C_1 symmetry respectively. The tabulated aqueous heats of formation, errors and details of the imaginary frequencies are shown in Table 5.1. Ball and stick diagrams for the reactants and products, and the transition states and intermediate structures for conformer C are shown in Figure 5.13. This shows how the reaction progresses for conformer C, and how the transition states differ from the intermediate structures during reaction.

Throughout the systems the solvated enthalpy of formation given by COSMO was lower in energy than the enthalpy of formation with Gibbs free energy of solvation given by SM5.4. The 'symmetrical' intermediate conformers have differing energies with AM1/COSMO favouring structure B, by 5.0 kJ mol^{-1}, over structure C and AM1/SM5.4A favouring structure A, by 6.6 kJ mol^{-1}, over structure C. Both methods lowest energy structures were the more symmetrical structures with the second lowest energy being structure C. Therefore it is necessary to study all the

possible intermediates when investigating a reaction mechanism. This is shown in Reaction 3, Section 5.3.3, where the reaction profiles for each of the conformers are calculated with all of the available semiempirical continuum methods. These studies indicate that the symmetrical species alone may not provide all the information about the overall reaction mechanism.

The reaction energy profiles show that the formation of TS_1 takes a substantial amount of energy, 74 to 78 and 19 to 27 kJ mol^{-1} for AM1/COSMO and AM1/SM5.4A respectively. There is then a large release of energy upon formation of the Meisenheimer complex (MC) of 111 to 117 and 107 to 125 kJ mol^{-1} for AM1/COSMO and AM1/SM5.4A respectively. This combined with a 132 to 137 kJ mol^{-1} barrier for the formation of TS_2 for the 'non-symmetrical' structure C produces a deep energy well for all the profiles. All the structures and methods produced a stepwise reaction profile with a deep energy well for the formation of the intermediate.

The MC for all the conformers was characterised as an intermediate; although structure A showed a single spurious imaginary of 5*i* and 15*i* cm^{-1} for AM1/COSMO and AM1/SM5.4A and structure B showed three 35*i*, 31*i* and 21*i* cm^{-1} for AM1/COSMO. The MC is lower in energy than the combined reactants/products by 33 to 39 and 89 to 98 kJ mol^{-1} for AM1/COSMO and AM1/SM5.4A respectively. Due to the comparative stability of the intermediate over the reactants and products, it could possibly be isolated at low temperatures, like in the analogous Meisenheimer reaction[7].

It was not possible to obtain a TS with only one imaginary frequency for many of the structures, although in each case the imaginary frequency corresponding to the desired transition vector was much larger and more significant than the others. The significant imaginary frequencies shown by AM1/SM5.4A had values of 313*i* cm^{-1} (TS_1) and 258*i* cm^{-1} (TS_2) for structure A and 285*i* cm^{-1} for structure B and 294*i* cm^{-1} (TS_1) and 220*i* cm^{-1} (TS_2) for Structure C. The significant imaginary frequencies shown by AM1/COSMO had values of 448*i* cm^{-1} (TS_1) and 424*i* cm^{-1} (TS_2) for structure A, 343*i* cm^{-1} for structure B and 403*i* cm^{-1} (TS_1) and 419*i* cm^{-1} (TS_2) for structure C. With both computational methods the transition structures have either the attacking nucleophile or the leaving group at a distance of 2.03 to 2.09 Å from the substrate.

The gradient norms for the optimisation calculations were all under 4 kJ per reduced unit for both the AM1/SM5.4A and the AM1/COSMO reactant structures, and below 6 kJ mol^{-1} per reduced unit for all the AM1/SM5.4A structures. This indicates that the structures produced have been reliably optimised. The AM1/COSMO method produced high gradient norms for for structure A (TS_1) and (TS_2) and structure C intermediate, TS_1 and TS_2 of 280, 257, 32, 155 and 257 kJ mol^{-1} per reduced unit respectively. This demonstrates the error and instability of optimisations using the AM1/COSMO method and is the main reason for the TSs having spurious imaginary frequencies.

Experimentally this reaction was found to be stepwise, as expected for an activated benzene derivative, reacting via a classical MC type intermediate. Computationally the reaction was also found to be stepwise with a large energy well for the formation of the MC. Both computational methods therefore reproduced and agreed with experimental results in showing a stepwise reaction for the 3-nitrophenolysis of 3'-nitrophenyl-2,4,6-trinitrophenyl ether.

Table 5.1 - The AM1 heat of formation energies in kJ mol^{-1} (ΔH_f), error (gradient norms), number of imaginary frequencies (NIMAG) and the value of imaginary frequencies in wavenumbers (cm^{-1}) or the chemically significant structures found for reaction 1.

AM1		MOPAC COSMO	AMSOL SM5.4 Water
3-Nitrophenoate Anion	ΔH_f	-529.96	-447.33
Charge = -1	Gradient Norm	0.04	0.01
	NIMAG	0	0
C3N3(OC6H5)2Py	ΔH_f	80.17	244.48
Charge = Neutral	Gradient Norm	3.01	3.68
C2	NIMAG	0	0
Reactants Combined DHf	ΔH_f	-449.79	-202.86
C3N3(OC6H5)2(Py)2			
Structure A	ΔH_f	-483.20	-300.62
Intermediate structure	Gradient Norm	2.60	0.95
Charge = -1	NIMAG	1	1
C_1		5i	15i
Structure A	C-O bond lengths	2.04/1.38	2.09/1.38
Transition structure 1	ΔH_f	-371.50	-175.84
	Gradient Norm	257.27	5.56
	NIMAG	2	2
	Frequency (cm^{-1})	424i, 7i	258i, 6i
Structure A	C-O bond lengths	1.37/2.03	1.40/2.04
Transition structure 2	ΔH_f	-345.71	-166.43
	Gradient Norm	279.87	3.77
	NIMAG	4	2
	Frequency	448i,37i,29i,16i	313i,44i
Structure B	ΔH_f	-489.26	-291.49
Intermediate structure	Gradient Norm	7.15	0.31
Charge = -1	NIMAG	3	0
C_2	Frequency	35i, 31i, 21i	
Structure B	C-O bond lengths	2.04/1.38	2.04/1.38
Transition structures	ΔH_f	-375.53	-184.17
	Gradient Norm	3.64	2.49
	NIMAG	3	2
	Frequency	343i, 29i, 15i	285i, 26i
Structure C	ΔH_f	-484.34	-294.07
Intermediate structure	Gradient Norm	32.09	0.86
Charge = -1	NIMAG	0	0
C_1	Frequency		
Structure C	C-O bond lengths	2.05/1.38	2.08/1.38
Transition structure 1	ΔH_f	-373.09	-177.44
	Gradient Norm	154.81	2.97
	NIMAG	3	1
	Frequency	403i,23i,16i	294i
Structure C	C-O bond lengths	1.37/2.03	1.35/2.09
Transition structure 2	ΔH_f	-350.49	-162.13
	Gradient Norm	271.29	2.93
	NIMAG	2	1
	Frequency	419i,19i	220i

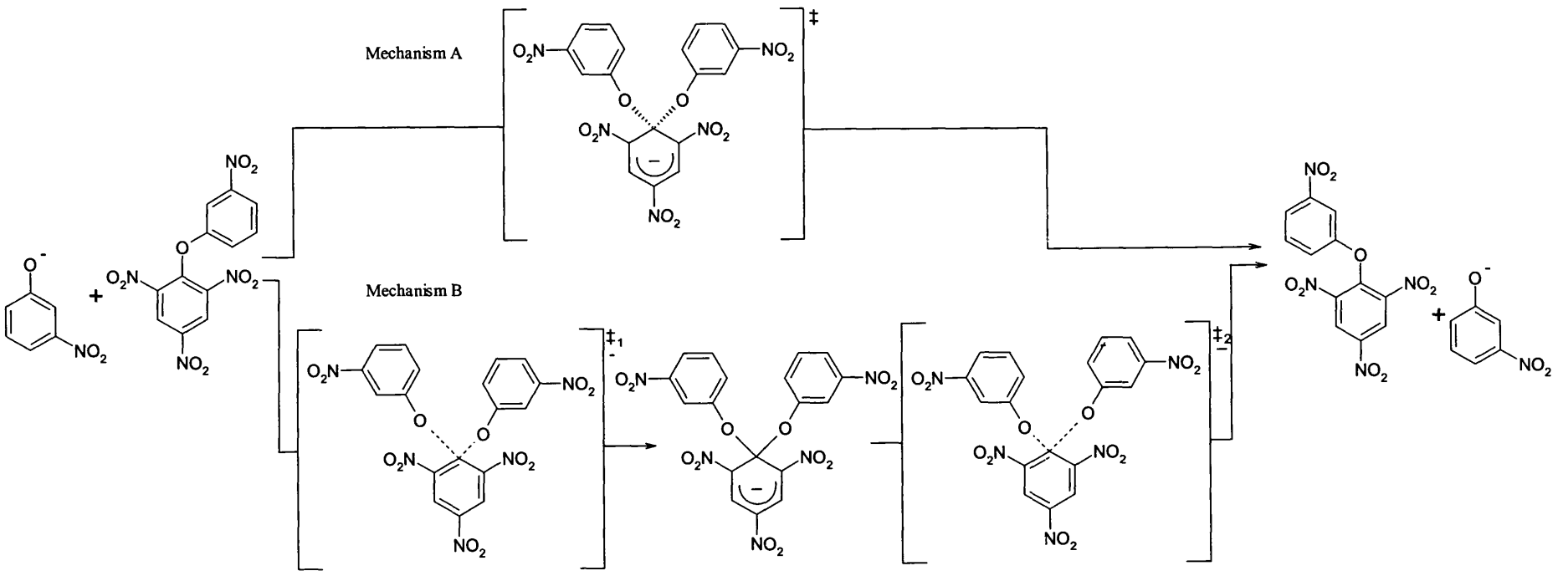
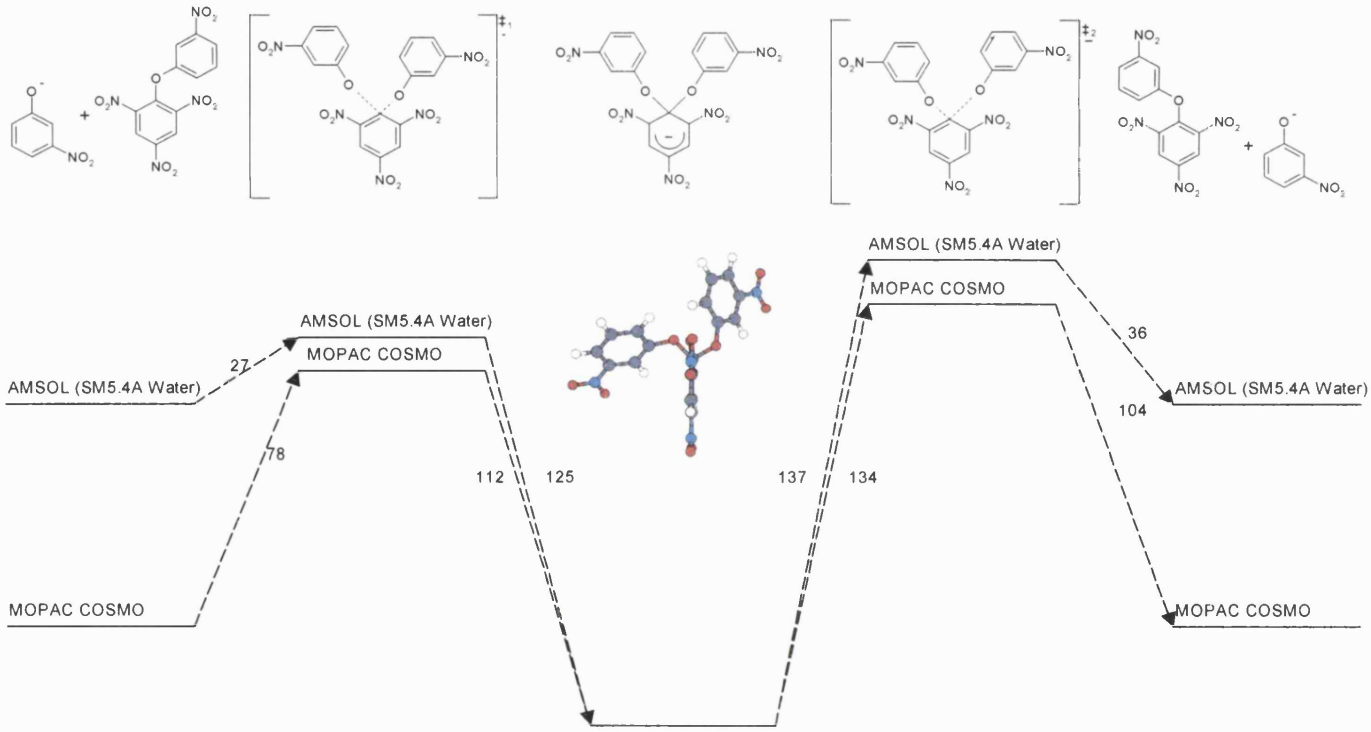
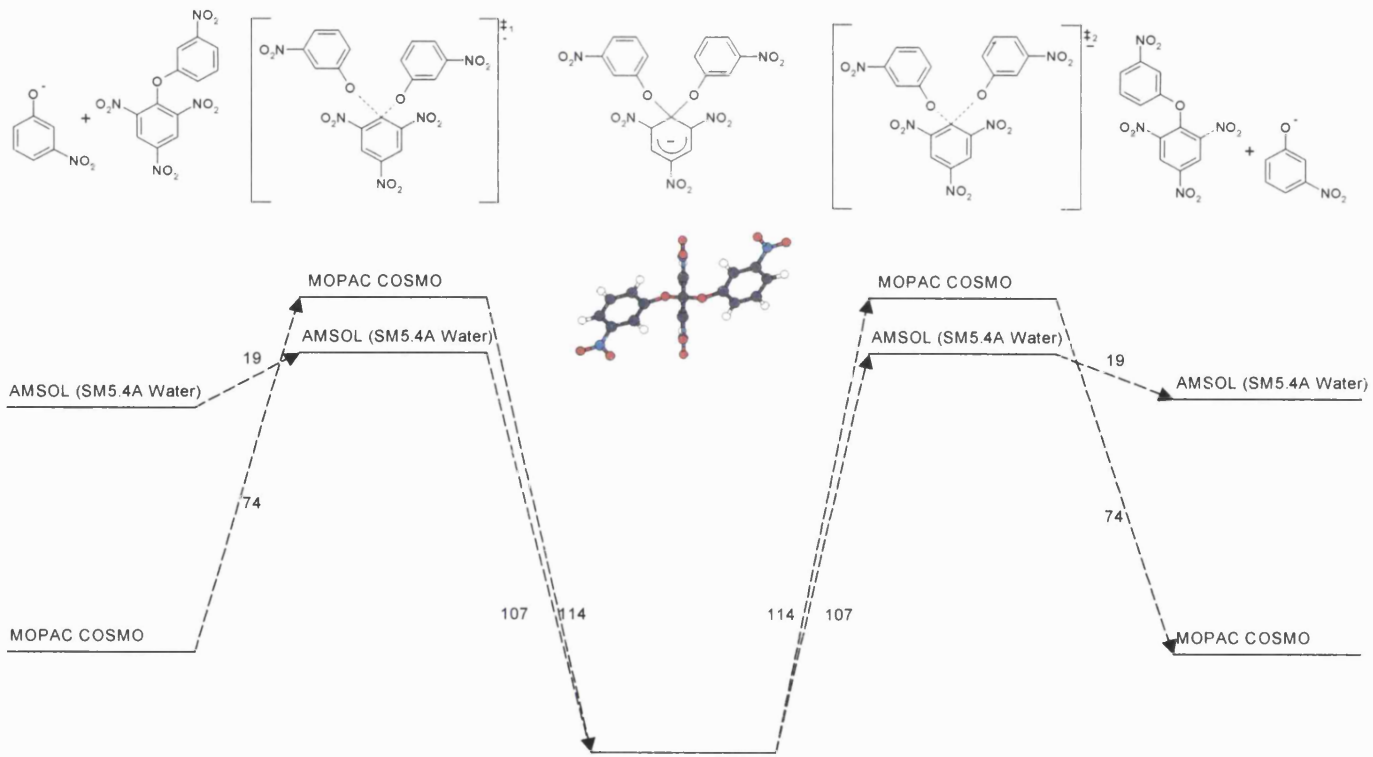


Figure 5.8 – the possible nucleophilic aromatic substitution mechanisms for the 3-nitrophenolysis of 3'-nitrophenyl-2,4,6-trinitrophenyl ether (Reaction 1). Mechanism A is the concerted (A_ND_N) mechanism and mechanism B is the stepwise (A_N + D_N) mechanism.



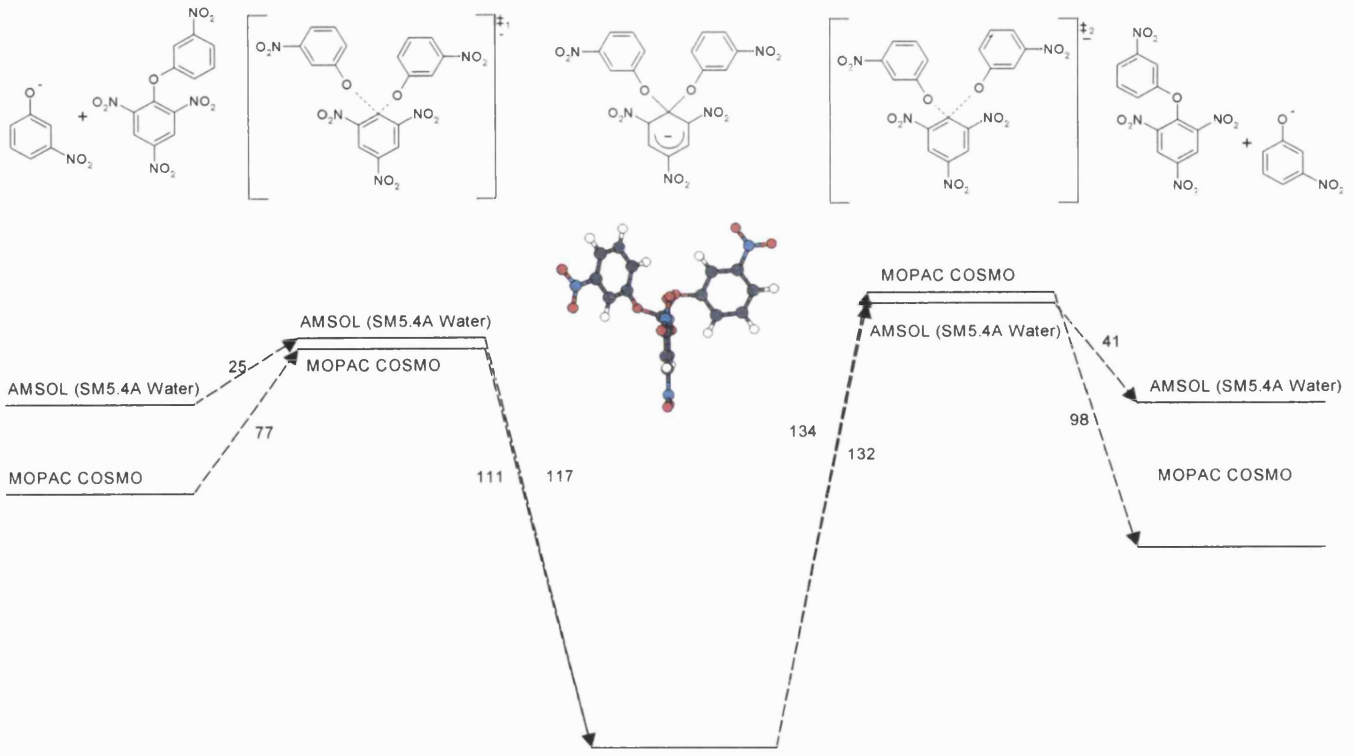
Structure A	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (COSMO)	-449.8	-345.7	-483.2	-371.5	-449.8
AMSOL (SM5.4 water)	-202.9	-166.4	-300.6	-175.8	-202.9

Figure 5.9 - The reaction energy profile relative to the aqueous energy (ΔH_f), in kJ mol^{-1}, of the intermediate for the 3-nitrophenoylsis of the 3'-Nitrophenyl 2,4,6-trinitrophenyl ether. (Structure A)



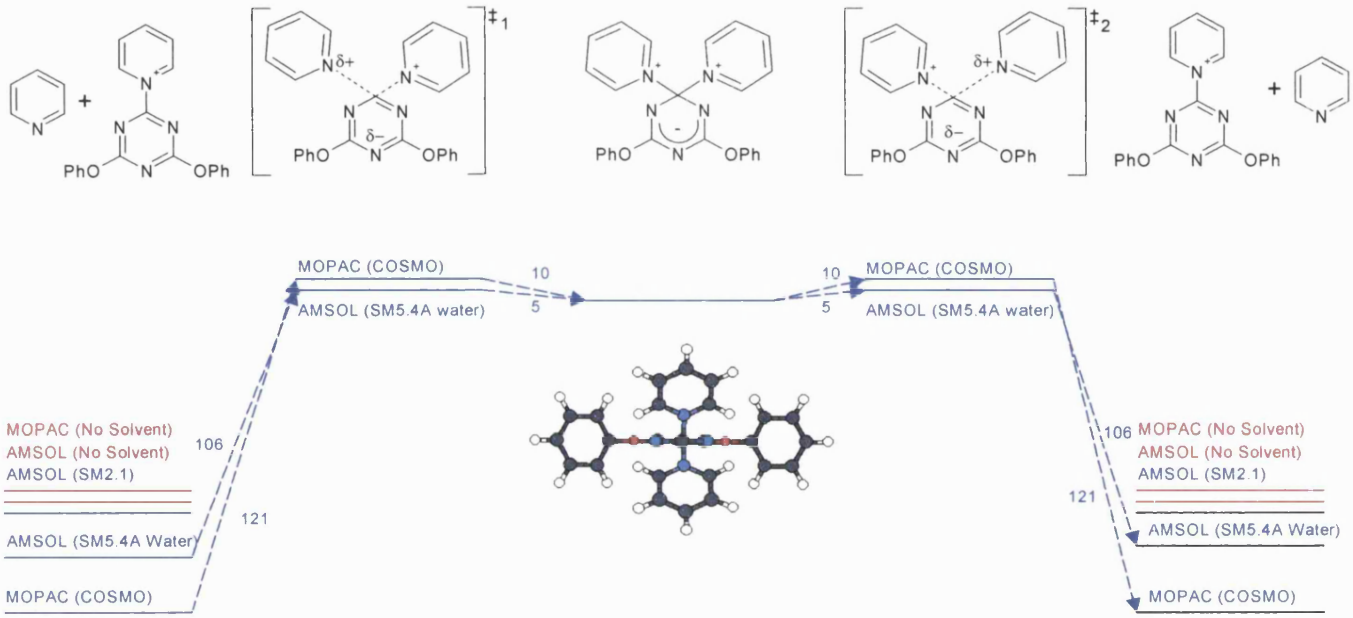
Structure B	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (COSMO)	-449.8	-375.5	-489.3	-375.5	-449.8
AMSOL (SM5.4 water)	-202.9	-184.2	-291.5	-184.2	-202.9

Figure 5.10 - The reaction energy profile relative to the aqueous energy (ΔH_f), in kJ mol^{-1}, of the intermediate for the 3-nitrophenoylsis of the 3'-Nitrophenyl 2,4,6-trinitrophenyl ether. (Structure B)



Structure C Computational method	Energy ΔH_f Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (COSMO)	-449.8	-373.1	-484.3	-350.5	-449.8
AMSOL (SM5.4 water)	-202.9	-177.5	-294.0	-162.1	-202.9

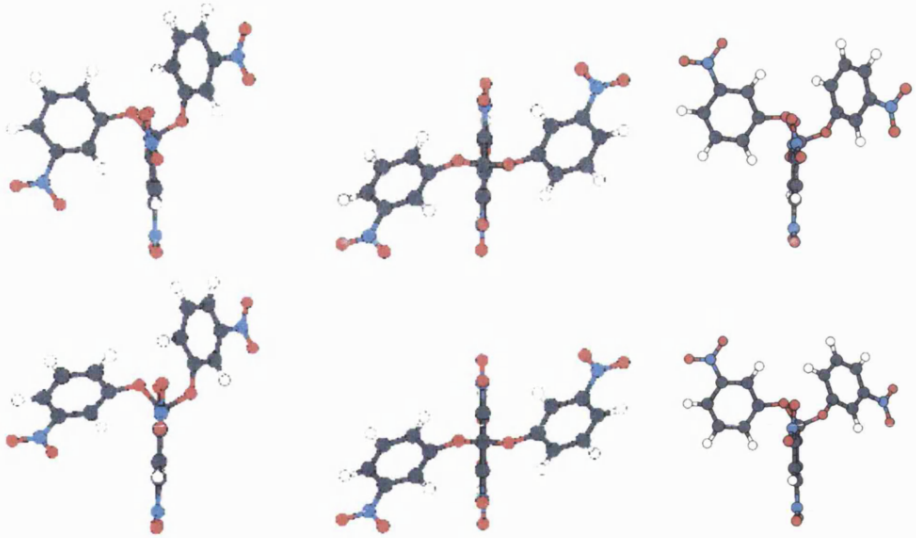
Figure 5.11 - The reaction energy profile relative to the aqueous energy (ΔH_f), in kJ mol^{-1}, of the intermediate for the 3-nitrophenoylsis of the 3'-Nitrophenyl 2,4,6-trinitrophenyl ether. (Structure C)



Computational method	Energy ΔH_f Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (No solvent)	1314.9		1420.9		1314.9
AMSOL (No solvent)	1316.5		1437.8		1316.5
MOPAC (COSMO)	1016.5	1162.2	1151.8	1162.2	1016.5
AMSOL (SM5.4 water)	1096.1	1231.1	1226.3	1231.1	1096.1
AMSOL (SM2.1 water)	1117.0		1243.9		1117.0

Figure 5.15 - The reaction profile relative to the heat of formation (AM1), in kJ mol^{-1}, of the intermediate for the pyridinolysis of the 1-(4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation

A (C_1) B (C_2) C (C_1)



Computational method	Intermediate Structure A	Intermediate Structure B	Intermediate Structure C
MOPAC (AM1/COSMO)	-483.2	**-489.3**	-484.3
AMSOL (AM1/SM5.4A)	**-300.6**	-291.5	-294.0

Figure 5.12 - Diagrams of the three possible structural isomers for the intermediate of the reaction of 3-nitrophenoylsis with 3'-Nitrophenyl-2,4,6-trinitrophenyl ether, as calculated by AM1/COSMO. The energies given are the solvated enthalpy of formation AM1/COSMO (bottom) and enthalpy of formation with Gibbs free energy of solvation AM1/SM5.4A (top), in kJ mol^{-1}.

$C_6H_4(NO_2)O^- + C_6H_2(NO_2)_3(OPh_2NO_2) \rightarrow C_6H_2(NO_2)_3(OPh_2NO_2)_2^- \rightarrow C_6H_4(NO_2)O^- + C_6H_2(NO_2)_3(OPh_2NO_2)$

Reactant/Product TS_1 Intermediate TS_2

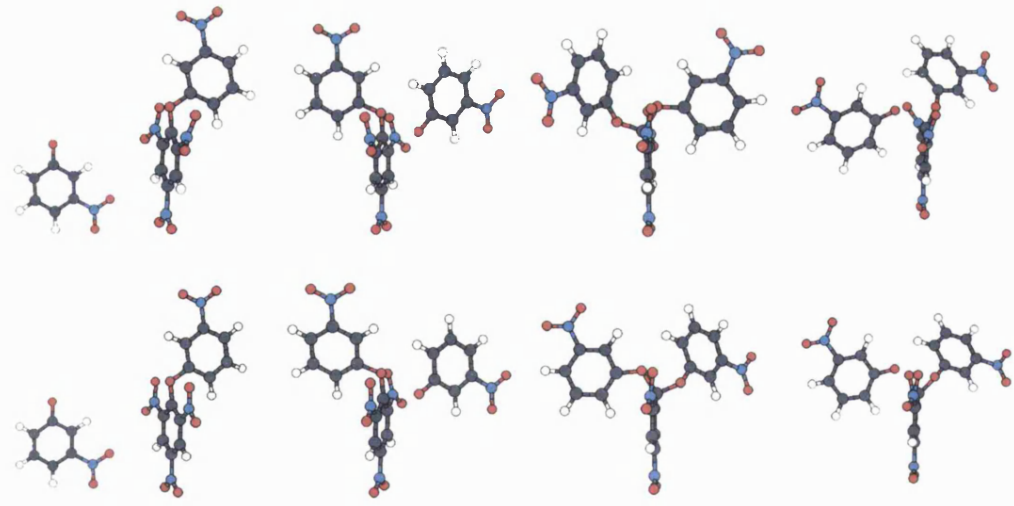


Figure 5.13 – Diagrams of the reactant and product structures and the intermediate and transition structures for conformer C for the 3-nitrophenoylsis of the 3'-Nitrophenyl-2,4,6-trinitrophenyl ether using the AM1/SM5.4A (top) and AM1/COSMO (bottom) methods.

5.2.2) The pyridinolysis of the 1-(4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation (Reaction 2: Experimentally stepwise)

Alternative nucleophilic reaction substitution mechanisms for Reaction 2 are shown in Figure 5.14. The reaction energy profile calculated by AM1 and PM3 methods are shown in Figures 5.15 and 5.16 respectively. The tabulated aqueous heats of formation, errors and details of the imaginary frequencies and ball and stick diagrams for the chemically significant structures found are shown in Table 5.2 and Figure 5.17 respectively.

This reaction was investigated using the AM1, PM3, AM1/COSMO, PM3/COSMO, AM1/SM5.4A and PM3/SM5.4P computational methods. The intermediate has two conformers, but since one has significantly higher energy, only the lower energy structure was focused upon. Unlike Reactions 1 and 3, which possess flexible aryloxy links to the substrate from the nucleophile and leaving group, the intermediate for Reaction 2 has less conformational freedom. The structure studied exhibits C_2 and C_1 symmetry; the low symmetry is due to skewing of the pyridine groups, particularly in the solvated calculations. The skewing may be caused by the formation of a partial hydrogen bond between a triazinyl nitrogen and one of the hydrogens on the pyridine ring.

Overall the AM1 energy profile, Figure 5.15, shows a stepwise mechanism with a small energy well. The energy difference between the reactant/product and intermediate structures is over 100 kJ mol^{-1} in both the gas and aqueous phase. In the aqueous phase the formation of TS_1 requires 106 and 121 kJ mol^{-1} for AM1/COSMO and AM1/SM5.4A respectively. The subsequent formation of the intermediate releases only a small amount of energy, under 10 kJ mol^{-1}.

The PM3 reaction energy profile, Figure 5.16, shows a stepwise aqueous energy profile with a significant energy well. A significant amount of energy is required to form TS_1, particularly in the aqueous phase with 79 (PM3/COSMO) and 57 kJ mol^{-1} (PM3/SM5.4P). There is then a release of energy of 33 (PM3/COSMO) and 22 kJ mol^{-1} (PM3/SM5.4P) markedly more than the AM1 aqueous calculations. The gas phase PM3 calculations have a comparatively small energy for formation of TS_1 of 18-25 kJ mol^{-1} and further energy is required in order to form the intermediate structure. This infers the reaction to proceed via a concerted

mechanism, however both the intermediate and TS have been characterised as such.

No TS could be found with AM1 using either AMSOL or MOPAC; this indicates a shallow energy well. The significant imaginary frequencies found for the TSs have been visualised and all occur along the reaction coordinate. The imaginary frequencies for AM1/COSMO, AM1/SM5.4A, PM3, PM3/COSMO, and PM3/SM5.4P were 406*i*, 319*i*, 144-146*i*, 313*i* and 268*i* cm^{-1} respectively. These are similar to expectations, however the gas phase PM3 method produced low imaginary frequencies. Extra spurious minor imaginary frequencies were found, number of extra imaginary frequencies in brackets, for the AMSOL AM1 intermediate (1), MOPAC PM3 intermediate (1), MOPAC PM3 TS (1), PM3/COSMO intermediate (1) and PM3/COSMO TS (2). This is partly caused by the high errors and instability of the optimisation procedures, shown in the high gradient norms, particularly for the PM3/COSMO method. The bond lengths to the substrate from the nucleophile (TS_1) or leaving group (TS_2) for AM1/COSMO, AM1/SM5.4A, PM3, PM3/COSMO, and PM3/SM5.4P were 1.81, 1.75, 1.85, 1.98 and 1.97 Å respectively. The aqueous PM3 calculations had the longest, nearly 2 Å bond lengths, and the AM1/SM5.4A method had the shortest.

Both AMSOL and MOPAC use different optimisation procedures and therefore both sets of gas phase calculations are presented here for completeness. (Gas phase calculations are approximately an order of magnitude less computationally intensive than the equivalent aqueous calculations). By comparing the energies it is clear the energies vary, depending on the program and therefore the optimisation procedure used. The change in the energy is generally in the first decimal place; this shows the limitations in precision of both the optimisation procedures and the methods.

The gas and aqueous phase AMSOL calculations had a lower gradient norm than MOPAC, indicating that AMSOL contains a more robust optimisation procedure. All of the aqueous AMSOL/SM5.4*x* structures having a gradient norm below 10 kJ per reduced unit, whereas the MOPAC/COSMO had a high gradient norm for the AM1 intermediate, PM3 intermediate and PM3 TS of 108, 99 and 85 kJ per reduced unit respectively.

Both experimental and computational methods showed a stepwise mechanism. The aqueous phase calculations have shallow well depths for the aqueous calculations of only 5, 10, 18 and 25 kJ mol^{-1} for AM1/SM5.4, AM1/COSMO, PM3/SM5.4P and PM3/COSMO respectively. The aqueous AM1 well depths were within the errors quoted for the solvation methods, adding doubt to the reliability of the computationally predicted mechanism.

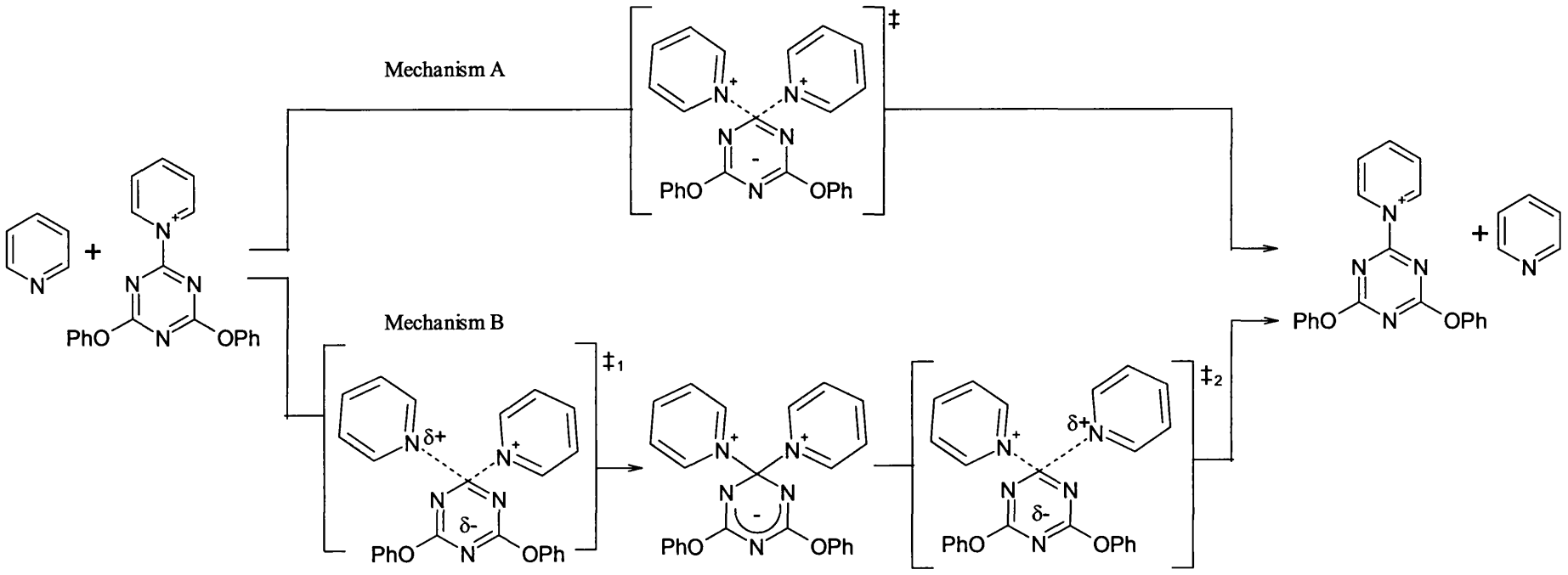
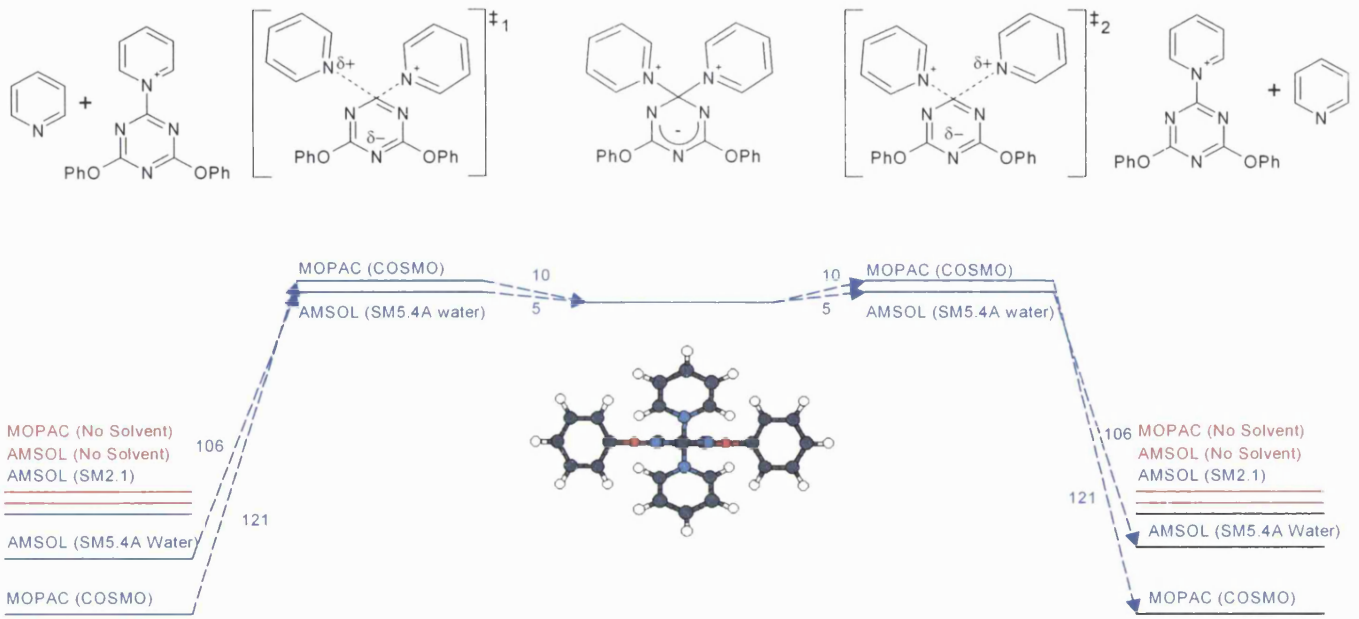


Figure 5.14 – the possible nucleophilic aromatic substitution mechanisms for the pyridinolysis of the (4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation (Reaction 2). Mechanism A is the concerted (A_ND_N) mechanism and mechanism B is the stepwise (A_N + D_N) mechanism.

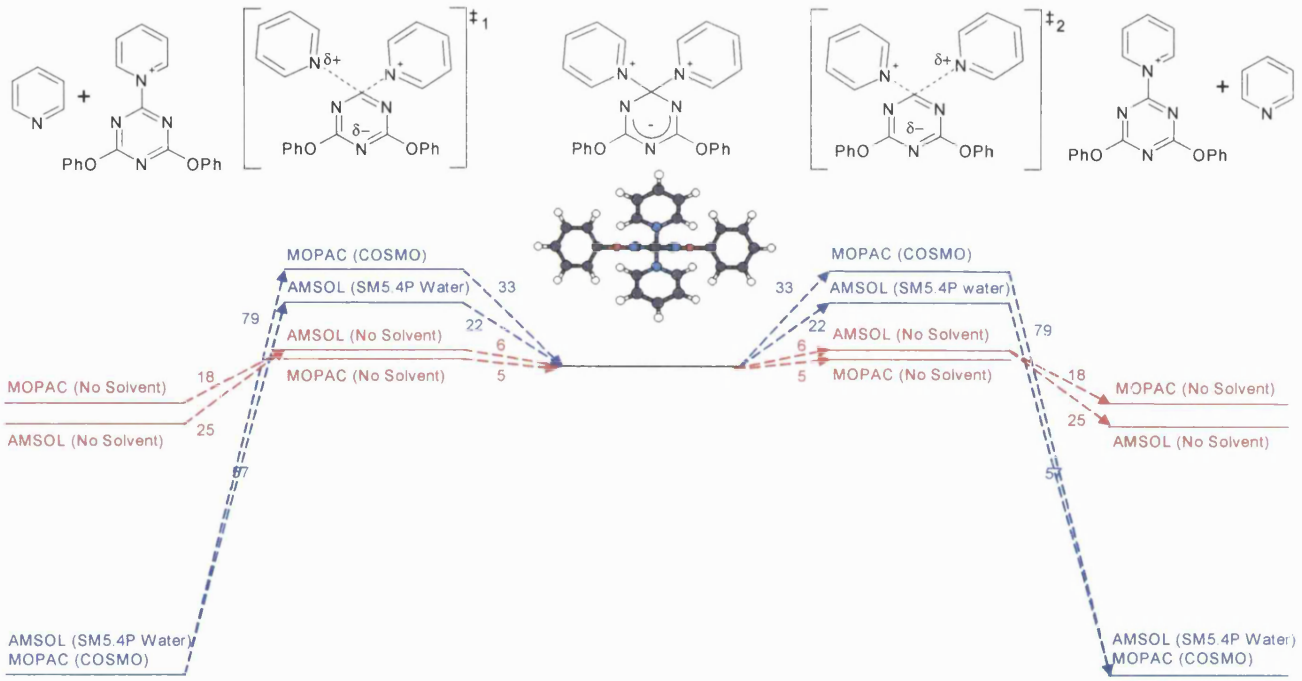
Table 5.2 - The AM1 and PM3 heat of formation energies (ΔH_f) in kJ mol^{-1}, errors (gradient norms in kJ per reduced unit), number of imaginary frequencies (NIMAG) and the value of imaginary frequencies in wavenumbers (cm^{-1}) for the chemically significant structures found for Reaction 2.

		MOPAC		AMSOL		
		No Solvent	COSMO	No Solvent	SM5.4 Water	SM2.1
AM1						
Pyridine	ΔH_f	133.78	101.35	134.05	109.18	115.30
Charge = Neutral	Gradient Norm	5.42	0.22	0.03	0.02	2.86
	NIMAG	0	0	0	0	
$C_3N_3(OC_6H_5)_2Py$	ΔH_f	1181.08	915.18	1182.43	986.90	1001.71
Charge = +1	Gradient Norm	0.03	50.44	0.04	0.10	4.39
C_2	NIMAG	0	0	0	0	
Reactants Combined	ΔH_f	1314.86	1016.53	1316.48	1096.08	1117.01
$C_3N_3(OC_6H_5)_2(Py)_2$						
Intermediate structur	ΔH_f	1420.89	1151.83	1437.79	1226.34	1244.10
Charge = +1	Gradient Norm	0.26	107.50	0.06	1.20	3.98
C_2	NIMAG	0	0	1	0	
	Frequency			21i		
Transition structure	C-N bond lengh	NOT	1.51/1.81	NOT	1.52/1.75	
	ΔH_f	FOUND	1162.24	FOUND	1231.07	
	Gradient Norm		2.25		5.67	
	NIMAG		1		1	
	Frequency		406i		313i	
PM3						
Pyridine	ΔH_f	133.76	101.34	127.07	106.73	
Charge = Neutral	Gradient Norm	0.54	0.12	0.02	0.01	
	NIMAG	0	0	0	0	
$C_3N_3(OC_6H_5)_2Py$	ΔH_f	939.68	680.15	940.98	749.15	
Charge = +1	Gradient Norm	2.64	7.36	0.03	0.09	
C_2	NIMAG	0	0	0	0	
Reactants Combined	ΔH_f	1073.45	781.49	1068.05	855.88	
$C_3N_3(OC_6H_5)_2(Py)_2$						
Intermediate structur	ΔH_f	1086.50	827.13	1088.20	891.53	
Charge = +1	Gradient Norm	2.55	98.70	0.02	0.08	
C_2	NIMAG	0	1	0	0	
	Frequency		24i			
Transition structure	C-N bond lengh	1.52/1.85	1.49/1.98	1.52/1.85	1.49/1.97	
	ΔH_f	1091.65	860.27	1093.28	913.16	
	Gradient Norm	1.92	84.98	0.05	1.42	
	NIMAG	2	3	1	1	
	Frequency	144i,28i	319i,74i,13i	146i	268i	



Computational method	Energy ΔH_f Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (No solvent)	1314.9		1420.9		1314.9
AMSOL (No solvent)	1316.5		1437.8		1316.5
MOPAC (COSMO)	1016.5	1162.2	1151.8	1162.2	1016.5
AMSOL (SM5.4 water)	1096.1	1231.1	1226.3	1231.1	1096.1
AMSOL (SM2.1 water)	1117.0		1243.9		1117.0

Figure 5.15 - The reaction profile relative to the heat of formation (AM1), in kJ mol^{-1}, of the intermediate for the pyridinolysis of the 1-(4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation



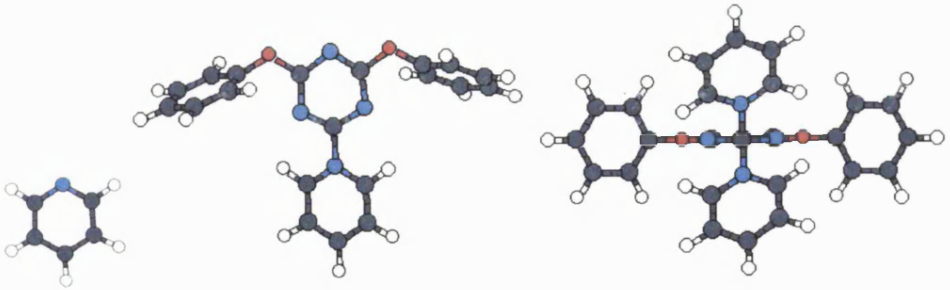
Computational method	Energy ΔH_f Reactants	Transition State 1	Intermediate	Transition State 2	Products
PM3					
MOPAC (No solvent)	1073.4	1091.6	1086.5	1091.6	1073.4
AMSOL (No solvent)	1068.0	1093.3	1088.2	1093.3	1068.0
MOPAC (COSMO)	781.5	860.3	827.1	860.3	781.5
AMSOL (SM5.4 water)	855.9	913.2	891.5	913.2	855.9

Figure 5.16 - The reaction profile relative to the heat of formation (PM3), in kJ mol^{-1}, of the intermediate for the pyridinolysis of the 1-(4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation

$C_5NH_5 + C_3N_3(OPh_2)(C_5NH_5)^+ \rightarrow C_3N_3(OPh_2)(C_5NH_5)_2^+ \rightarrow C_5NH_5 + C_3N_3(OPh_2)(C_5NH_5)^+$

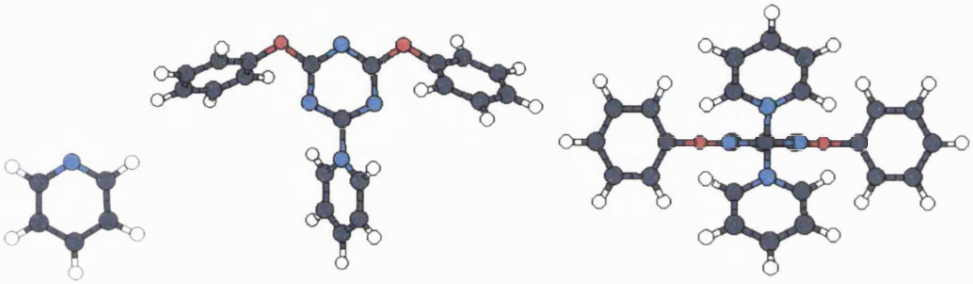
AMSOL (AM1)

Reactant/Product structures 'Intermediate' structure



AMSOL (AM1/SM5.4A water)

Reactant/Product structures 'Intermediate' structure



MOPAC (AM1/COSMO)

Reactant/Product structures 'Intermediate' structure

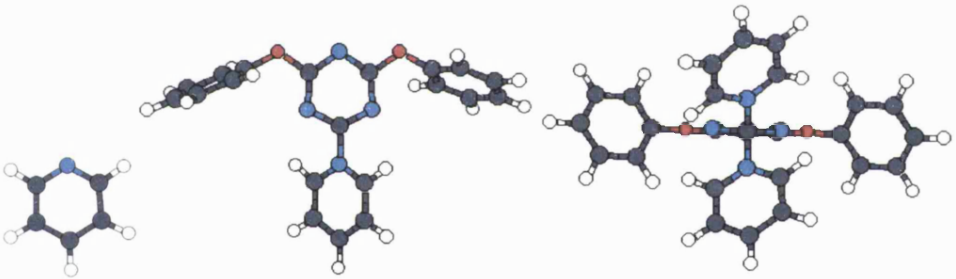


Figure 5.17 – Diagrams of the reactant, product and intermediate structures the pyridinolysis of the (4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation as calculated using the AM1 (top), AM1/COSMO (middle) and AM1/SM5.4A (bottom) methods.

5.2.3) The nucleophilic aromatic substitution reaction of 4-nitrophenolate anions with 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Reaction 3: Experimentally concerted)

Alternative nucleophilic reaction substitution mechanisms for Reaction 3 are shown in Figure 5.18. This reaction was investigated using the AM1, PM3, AM1/COSMO, PM3/COSMO, AM1/SM5.4A and PM3/SM5.4P computational methods. The tabulated aqueous heats of formation, errors and details of the imaginary frequencies for AM1 and PM3 are shown in Tables 5.3 and 5.4 respectively. There are three reaction profiles for each method because the "symmetrical" intermediate has three conformers due to the flexibility of the phenoxy link to the substrate, as in Reaction 1. Diagrams of the three intermediates structures, from the gas phase AM1 calculations, and the energies for all the methods are shown in Figure 5.25. The structures A, B and C have symmetry of C_2, C_{2v} and C_1 respectively. Ball and stick diagrams for the chemically significant reactant/product and intermediate structures using a selection of the methods (AM1, AM1/COSMO and AM1/SM5.4A methods) are shown in Figure 5.26.

The least symmetrical and yet least sterically hindered species of the conformers, structure C, is generally the lowest energy, otherwise it is the second lowest energy structure. The exceptions were the gas phase AM1 calculation, which favoured structure A by 2.7 kJ mol^{-1} and the PM3/COSMO calculations, which favoured structure B by 1.5 kJ mol^{-1} and PM3/SM5.4P, for which structure B and C had the same energy. The other methods favoured structure C by a larger energy margin. The steric effects of the flexible and bulky nucleophile and leaving group moieties could be significant as the least sterically hindered structure, structure C, is also the lowest energy. The symmetrical structure is therefore not the only structure to be focused upon, especially when flexible bulky groups are involved. This is computationally inconvenient, as the calculation of symmetrical structures require substantially less computational expense.

Both AMSOL and MOPAC gas phase calculations are presented here in order to show the effect of the different optimisers in these programs used. Given an identical structure, a single point calculation, both methods give almost identical energies differences occur due to slightly different parameters; for example, the

enthalpy calculated for phenol using AM1 was –81.640 and –81.301 kJ mol^{-1} for MOPAC 93 and AMSOL 6.1 respectively. MOPAC produced lower energy optimised structures than AMSOL for AM1 and all bar one of the PM3 calculations. The gas phase AMSOL calculations do however have lower gradient norms than the equivalent MOPAC calculations.

The intermediate structures have been characterised and have all real frequencies and each TS have a single significant imaginary frequency. The AM1 methods produced significant imaginary frequencies of 303*i* to 353*i* cm^{-1} and 422*i* and 480*i* cm^{-1} for the gas and aqueous phase TSs respectively. The PM3 results produced low significant imaginary frequencies for the TSs with values of 123*i* to 256*i* cm^{-1} and 299*i* to 441*i* cm^{-1} for the gas and aqueous phase TSs respectively. All the significant imaginary frequencies were along the reaction coordinate.

The gradient norms were all low (error in the optimisation procedure) under 5 kJ per reduced unit, for all of the calculations except for the COSMO method. The gradient norm for the COSMO calculations did not have a value below 5 kJ per reduced unit for any of the intermediate or TS structures, therefore, the structures found may not be sufficiently refined. There were a number of spurious minor imaginary frequencies (all under 100*i* cm^{-1}) particularly for the COSMO method, as outlined below. The AM1/COSMO method produced extra spurious minor imaginary frequencies, in brackets, for structure B intermediate (1), structure C intermediate (1) and TS_1 (1). The PM3/COSMO method produced extra spurious minor imaginary frequencies, in brackets, these were the intermediate A (2), B (1), C (1) and TS A (1), B (2) and C (2,2). The high gradient norms and the amount of spurious imaginary frequencies indicates that the optimisation procedure within MOPAC is not stable for calculating COSMO structures.

The AM1 reaction energy profiles for structures A, B and C are shown in Figures 5.19, 5.20 and 5.21. The PM3 reaction energy profiles for structures A, B and C are shown in Figures 5.22, 5.23 and 5.24. These profiles shall be discussed individually in turn. The AM1 energy profile of the structure A (Figure 5.19) shows a stepwise mechanism with a 22-23 kJ mol^{-1} energy well. The gas phase calculations show much lower activation energy required for the formation of TS_1, than the aqueous phase calculations at 15 versus 82 and 114 kJ mol^{-1} for COSMO and SM5.4A respectively. Structure B (Figure 5.20), like structure A, all show a

stepwise mechanism with a 31-34 kJ mol^{-1} energy well. The reaction profiles follow almost the same path with a smaller gas phase activation energy of 38 compared to 88 and 115 kJ mol^{-1} for the aqueous COSMO and SM5.4A respectively. Like the other conformers structure C (Figure 5.21) shows a stepwise mechanism with a 22-33 kJ mol^{-1} energy well. This non-symmetrical mechanism has differences for both steps, notably for COSMO method. The formation of two TSs is due to the non-symmetrical nature of the intermediate where the substrate, attacking and leaving group are arranged in a 'propeller' type arrangement. One group is perpendicular to the triazine ring and the second group skewed away from both the first group and the triazine ring therefore these bind and leave differently. The formation of TS_1 and the decomposition of TS_2 have a much larger change in energy for the aqueous phase than the gas phase. The PM3 energy profile of the structure A (Figure 5.22), like the AM1 calculations, shows a stepwise mechanism. The formation of the TS required 13 kJ mol^{-1} in the gas phase and 69 kJ mol^{-1} and 79 kJ mol^{-1} for COSMO and SM5.4P respectively. These had well depths of 16, 50 and 55 kJ mol^{-1} for gas phase, COSMO and SM5.4P respectively. The energy profile for structure B (Figure 5.23) follows a similar stepwise mechanism. The energy required for formation of TS_1 was markedly smaller for the gas phase, 18 kJ mol^{-1}, than the aqueous phase with COSMO and SM5.4P requiring 107 and 77 kJ mol^{-1} respectively. The energy well for the formation of the intermediate is similar for all of the methods at 56, 61 and 70 kJ mol^{-1} the gas phase, COSMO and SM5.4P methods respectively. The non-symmetrical structure, structure C, (Figure 5.25) follows a stepwise mechanism with a significant energy well, varying between 61 and 74 kJ mol^{-1}, shown for all the methods used. The two gas phase methods produced varying energy profiles, due to the optimisation methods used. This is notably a differing energy required to form of TS_1 18 and 30 kJ mol^{-1} for AMSOL and MOPAC respectively; the subsequent decomposition from TS_2 to products was 18 kJ mol^{-1} for both the methods. The aqueous phase formation of TS_1 and decomposition of TS_2 had the same energy change of 113 kJ mol^{-1} for COSMO and 90 and 77 kJ mol^{-1} for SM5.4P water respectively. All the AM1 and PM3 profiles have smaller activation and decomposition energies in the gas phase than in the aqueous phase calculations.

Structure C has lower total energies for the AM1 aqueous calculations and structures B and C have the lowest energy for the PM3 aqueous calculations.

Therefore, these are the most probable reaction paths with which experimental results can be compared. All of the methods and conformers investigated predicted a stepwise mechanism, this contradicts the experimental observation of a concerted mechanism.

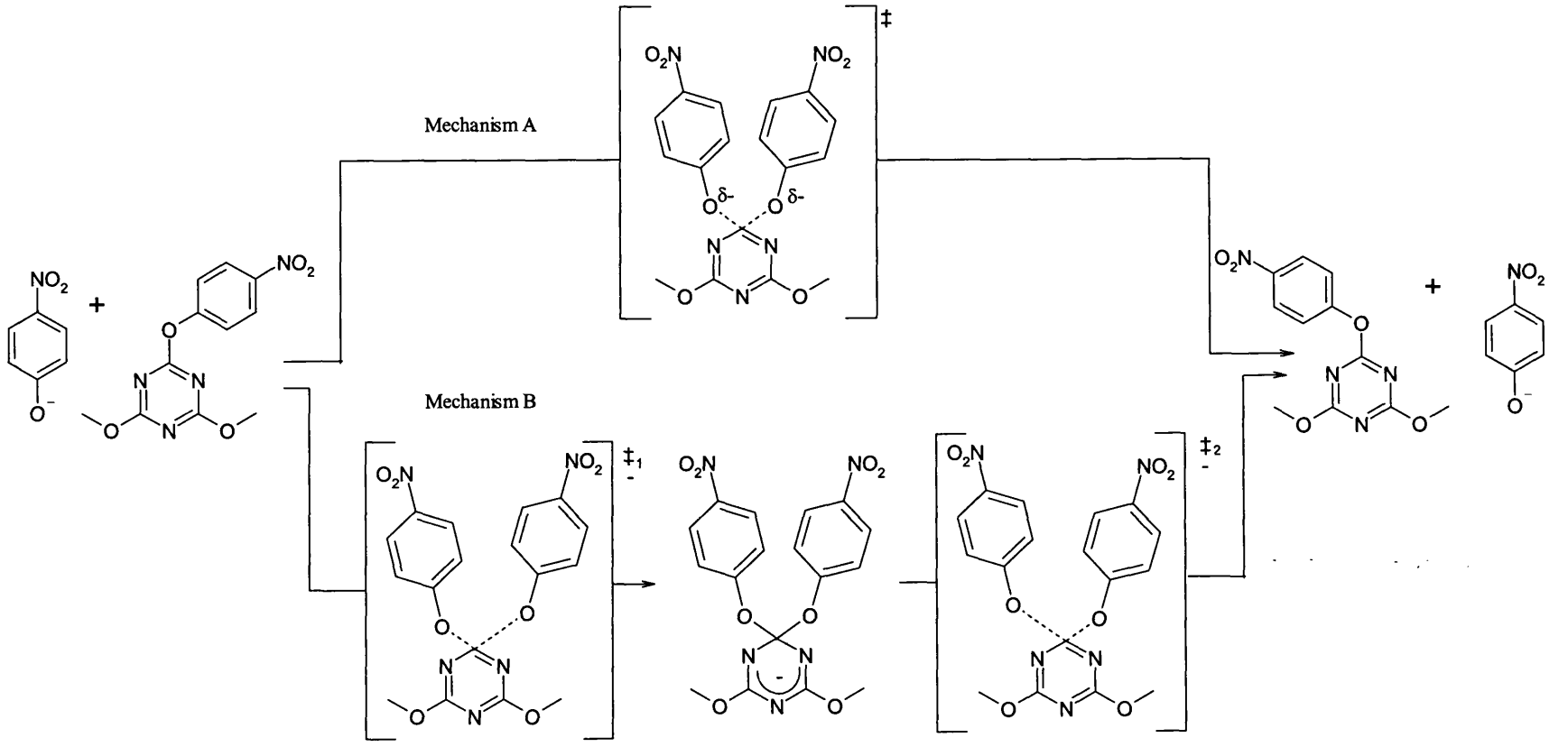


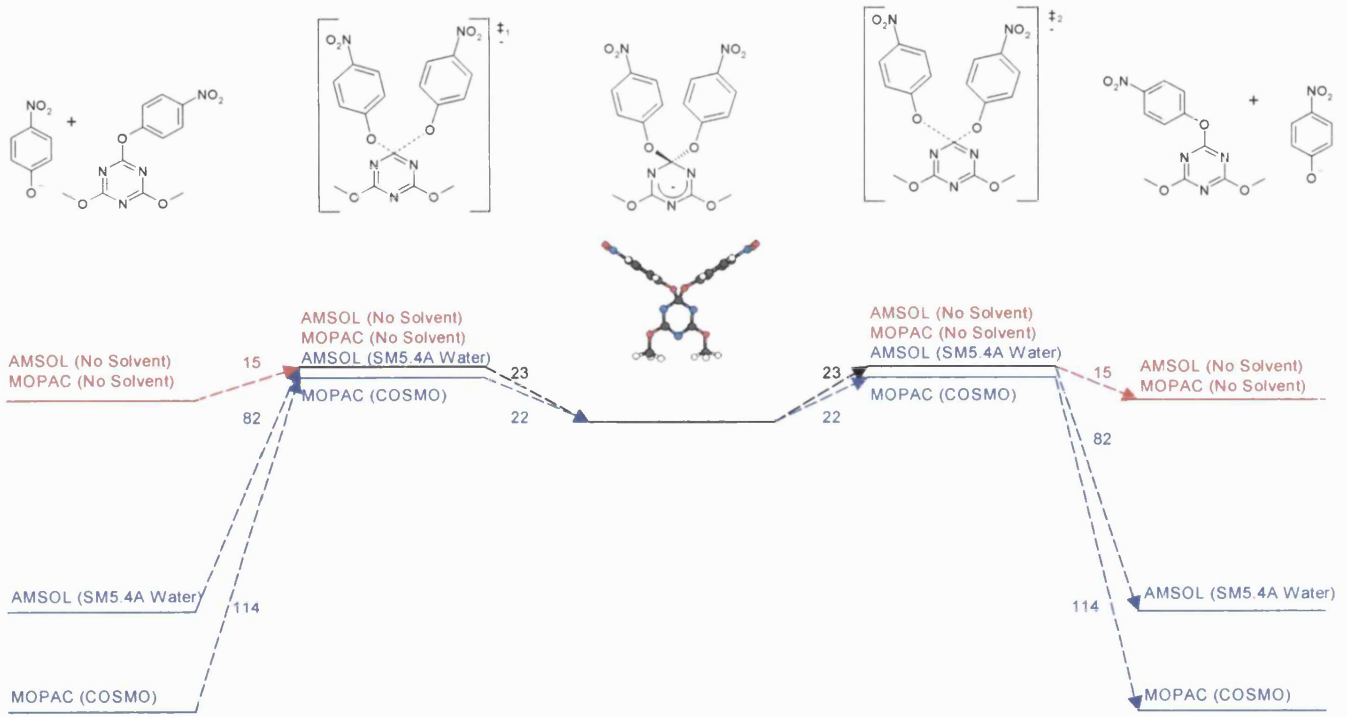
Figure 5.18 – the possible nucleophilic aromatic substitution mechanisms for the 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Reaction 3). Mechanism A is the concerted (A_ND_N) mechanism and mechanism B is the stepwise (A_N + D_N) mechanism.

Table 5.3 - The AM1 heat of formation energies in kJ mol^{-1} (ΔH_f), errors (gradient norms in kJ per reduced unit), number of imaginary frequencies (NIMAG) and the value of imaginary frequencies in wavenumbers (cm^{-1}) for the chemically significant structures found for Reaction 3. The SM2.1 results were calculated by Mr Philip Hoskin during his time in Bath.

AM1		MOPAC		AMSOL		
		No Solvent	COSMO	No Solvent	SM5.4 Water	SM2.1
$C_6H_4(NO_2)O$	ΔH_f	-272.47	-562.23	-271.95	-483.72	-494.12
Charge = -1	Gradient Norm	0.01	50.96	0.05	0.06	1.27
	NIMAG	0	0	0	0	
$C_3N_3(OMe)_2OPhNO_2$	ΔH_f	14.06	-110.50	15.19	-32.34	-34.16
Charge = Neutral	Gradient Norm	1.48	55.08	0.03	0.02	4.01
	NIMAG	0	0	0	0	
Reactants Combined ΔH_f		-258.42	-672.73	-256.76	-516.06	-528.28
$C_3N_3(OMe)_2(OPhNO_2)_2$						
Structure A	ΔH_f	-266.24	-580.23	-264.55	-456.56	
Intermediate structure	Gradient Norm	0.04	72.45	0.83	0.03	
Charge = -1	NIMAG	0	0	0	0	
C_2						
Structure A	C-O bond lengths	1.82/1.42	1.80/1.41	1.82/1.42	1.79/1.40	
Transition structures	ΔH_f	-242.95	-558.92	-241.27	-433.86	
	Gradient Norm	2.45	34.30	0.06	0.03	
	NIMAG	1	1	1	1	
	Frequency (cm^{-1})	352i	478i	353i	453i	
Structure B	ΔH_f	-251.55	-591.34	-249.87	-462.12	
Intermediate structure	Gradient Norm	15.01	46.97	0.11	0.11	
Charge = -1	NIMAG	0	1	0	0	
C_{2v}	Frequency		14i			
Structure B	C-O bond lengths	1.86/1.43	1.81/1.43	1.43/1.86	1.83/1.43	
Transition structures	ΔH_f	-220.47	-558.11	-218.82	-427.84	
	Gradient Norm	2.97	18.66	0.05	0.06	
	NIMAG	1	1	1	1	
	Frequency	322i	454i	326i	422i	
Structure C	ΔH_f	-264.78	-594.57	-263.09	-466.18	
Intermediate structure	Gradient Norm	0.02	124.42	0.02	0.14	
Charge = -1	NIMAG	0	1	0	0	
C_1	Frequency		17i			
Structure C	C-O bond lengths	1.87/1.42	1.42/1.82	1.87/1.42	1.82/1.42	
Transition structure 1	ΔH_f	-238.89	-566.46	-230.27	-433.76	
	Gradient Norm	2.26	4.69	0.04	0.10	
	NIMAG	1	2	1	1	
	Frequency	353i	457i,19i	303i	429i	
Structure C	C-O bond lengths	1.42/1.82	1.41/1.80	1.42/1.87	1.41/1.80	
Transition structure 2	ΔH_f	-242.94	-565.19	-237.20	-436.08	
	Gradient Norm	2.47	4.45	0.03	0.09	
	NIMAG	1	1	1	1	
	Frequency	352i	480i	352i	436i	

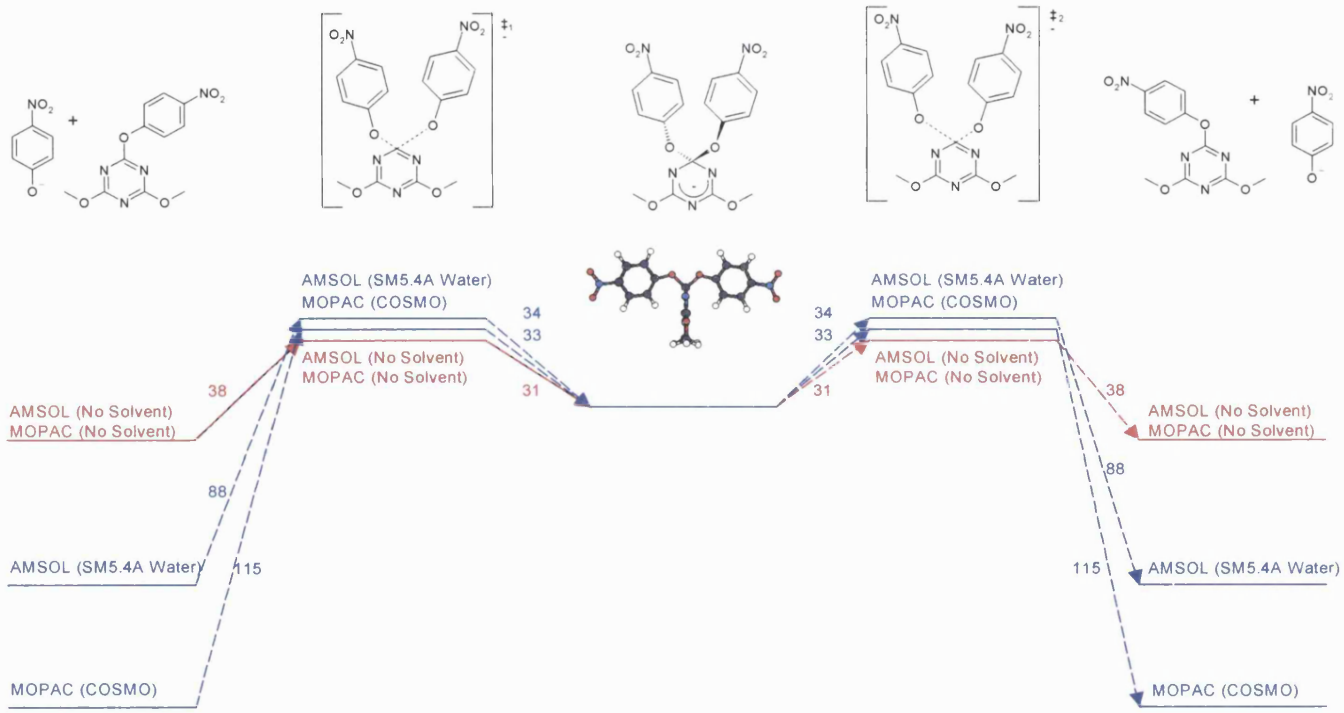
Table 5.4 - The PM3 heat of formation energies in kJ mol^{-1} (ΔH_f), errors (gradient norms in kJ per reduced unit), number of imaginary frequencies (NIMAG) and the value of imaginary frequencies in wavenumbers (cm^{-1}) for the chemically significant structures found for Reaction 3.

PM3		MOPAC		AMSOL	
		No Solvent	COSMO	No Solvent	SM5.4 Water
C_6H_4 (NO_2) O	ΔH_f	-341.67	-662.20	-341.18	-560.83
Charge = -1	Gradient Norm	0.98	20.00	0.07	0.46
	NIMAG	0	0	0	0
C_3N_3 $(OMe)_2$ $OPhNO_2$	ΔH_f	-221.53	-383.51	-220.51	-266.19
Charge = Neutral	Gradient Norm	2.30	3.77	0.02	0.03
	NIMAG	2	1	0	0
	Frequency	123i,88i	24i		
Reactants Combined ΔH_f		-563.21	-1045.71	-561.68	-827.02
C_3N_3 $(OMe)_2$ $(OPhNO_2)_2$					
Structure A	ΔH_f	-605.61	-991.98	-604.06	-797.54
Intermediate structure	Gradient Norm	3.53	6.28	0.02	5.29
Charge = -1	NIMAG	0	2	0	0
C_2	Frequency		46i,23i		
Structure A	C-O bond lengths	1.37/1.96	1.36/1.90	1.37/1.96	1.37/1.88
Transition structures	ΔH_f	-550.32	-976.38	-548.77	-747.59
	Gradient Norm	1.93	17.57	0.05	4.49
	NIMAG	1	2	1	1
	Frequency	239i	410i, 11i	240i	366i
Structure B	ΔH_f	-600.64	-1008.47	-599.37	-810.89
Intermediate structure	Gradient Norm	0.29	73.87	0.09	0.14
Charge = -1	NIMAG	0	1	0	0
C_{2v}	Frequency		24i		
Structure B	C-O bond lengths	1.973	1.90	1.95	1.90
Transition structures	ΔH_f	-544.96	-938.82	-543.42	-749.84
	Gradient Norm	2.66	38.74	0.08	2.14
	NIMAG	1	3	1	1
	Frequency	230i	412i,61i,4i	231i	300i
Structure C	ΔH_f	-607.12	-1005.84	-605.56	-810.93
Intermediate structure	Gradient Norm	1.89	12.01	0.02	2.05
Charge = -1	NIMAG	0	1	0	0
C_1	Frequency		25i		
Structure C	C-O bond lengths	1.3 9/2.04	1.39/1.94	1.36/1.97	1.39/1.87
Transition structure 1	ΔH_f	-533.67	-932.45	-543.41	-749.61
	Gradient Norm	1.42	18.66	0.05	0.71
	NIMAG	1	3	1	1
	Frequency	138i	358i,65i,32i	213i	368i
Structure C	C-O bond lengths	1.97/1.39	1.88/1.36	1.93/1.37	1.90/1.37
Transition structure 2	ΔH_f	-544.97	-932.45	-543.59	-749.86
	Gradient Norm	2.09	88.03	0.04	0.04
	NIMAG	1	3	1	1
	Frequency	231i	441i,51i,20i	256i	299i



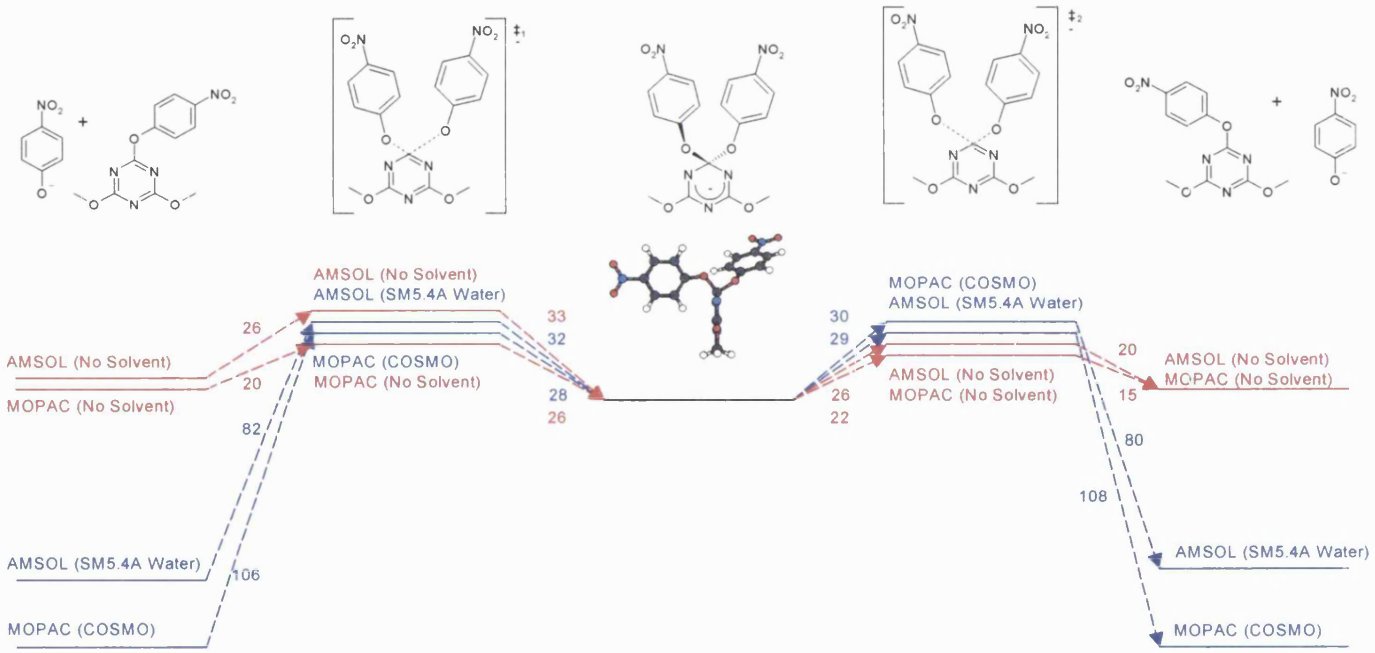
Structure A	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (No solvent)	-258.4	-242.9	-266.2	-242.9	-258.4
AMSOL (No solvent)	-256.8	-241.3	-264.6	-241.3	-256.8
MOPAC (COSMO)	-672.7	-558.5	-580.2	-558.5	-672.7
AMSOL (SM5.4 water)	-516.1	-433.8	-456.6	-433.8	-516.1

Figure 5.19 – Heats of formation (AM1), in kJ mol^{-1}, relative to the symmetrical intermediate for 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Structure A)



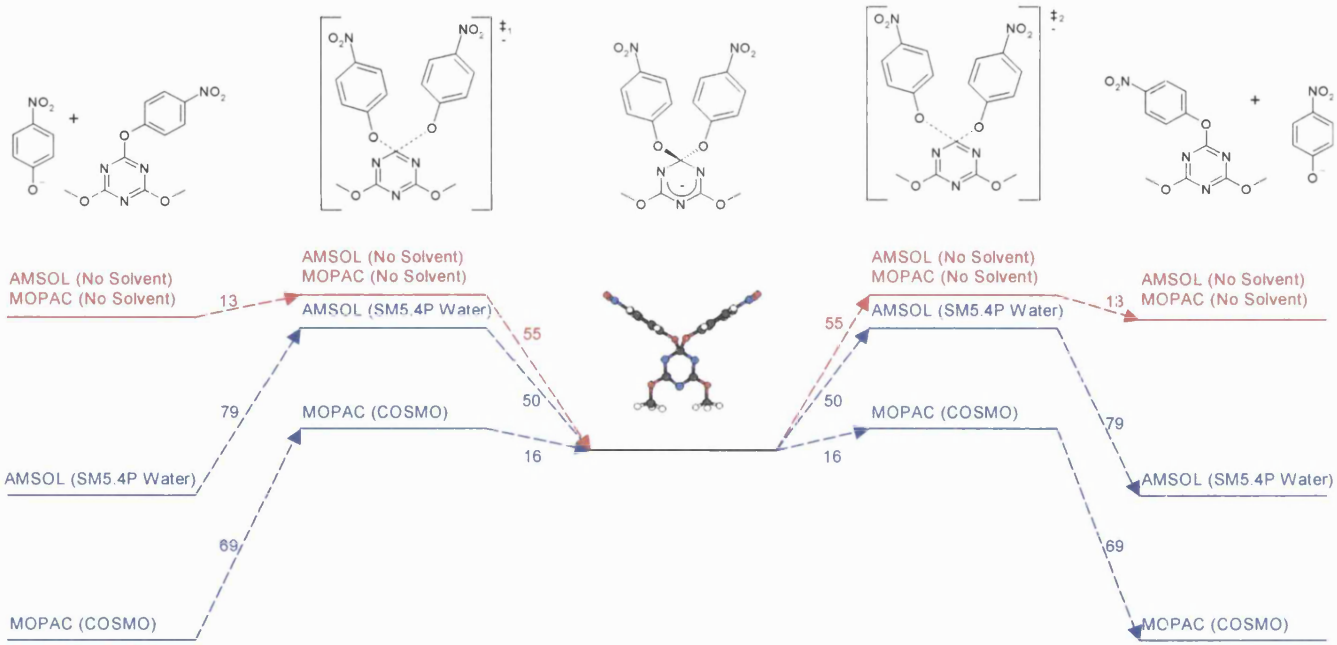
Structure B Computational method	Energy ΔH_f Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (No solvent)	-258.4	-220.5	-251.6	-220.5	-258.4
AMSOL (No solvent)	-256.8	-218.8	-249.9	-218.8	-256.8
MOPAC (COSMO)	-672.7	-558.1	-591.3	-558.1	-672.7
AMSOL (SM5.4 water)	-516.1	-427.9	-462.1	-427.9	-516.1

Figure 5.20 - Heats of formation (AM1), in kJ mol^{-1}, relative to the symmetrical intermediate for 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Structure B)



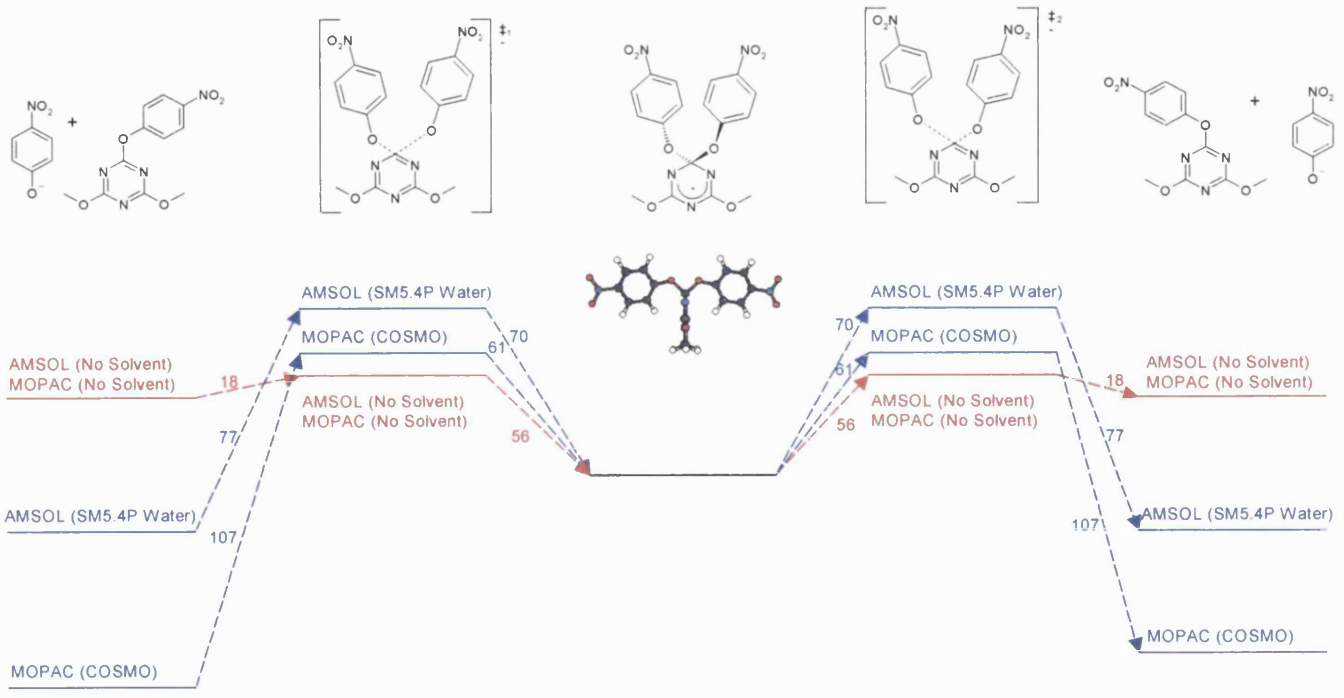
Structure C	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
AM1					
MOPAC (No solvent)	-258.4	-238.9	-264.8	-242.9	-258.4
AMSOL (No solvent)	-256.8	-230.3	-263.1	-237.2	-256.8
MOPAC (COSMO)	-672.7	-566.5	-594.6	-565.2	-672.7
AMSOL (SM5.4 water)	-516.1	-433.8	-466.2	-436.1	-516.1

Figure 5.21 - Heats of formation (AM1), in kJ mol^{-1}, relative to the intermediate for 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Structure C)



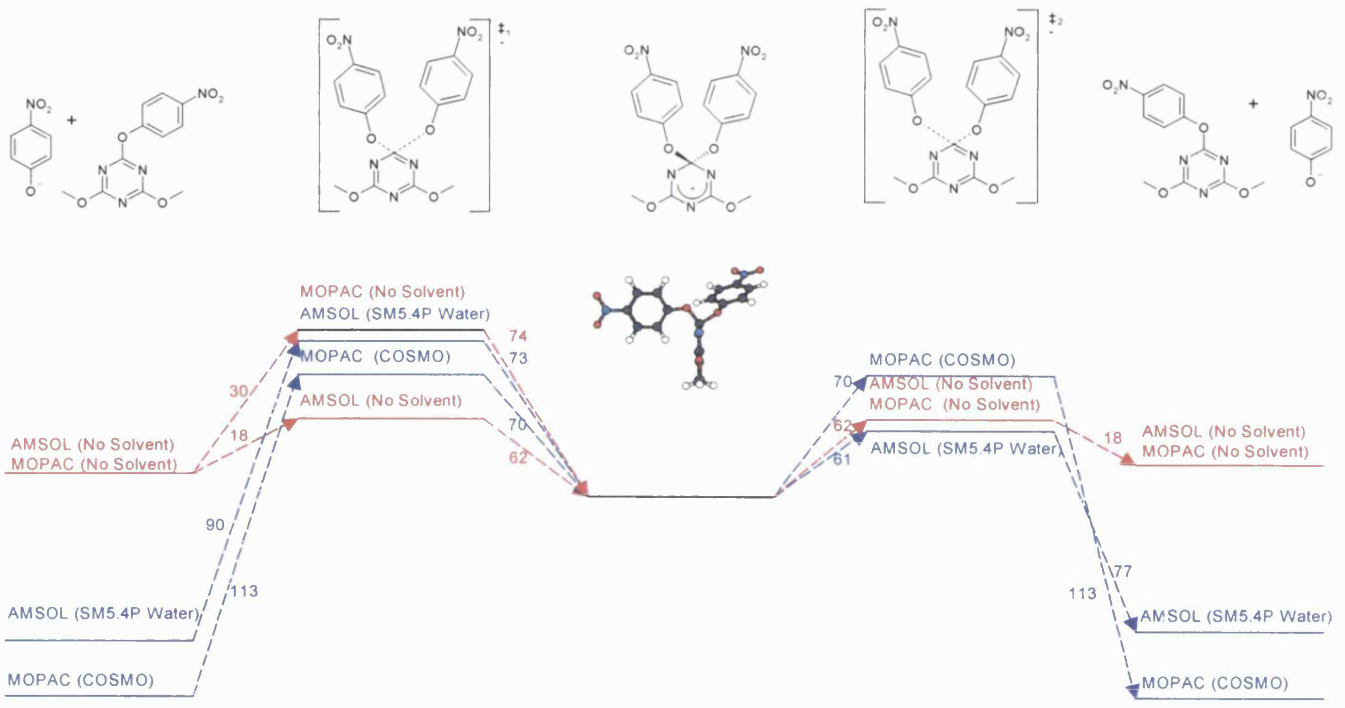
Structure A	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
PM3					
MOPAC (No solvent)	-563.2	-550.3	-605.6	-550.3	-563.2
AMSOL (No solvent)	-561.7	-548.8	-604.1	-548.8	-561.7
MOPAC (COSMO)	-1045.7	-976.4	-992.0	-976.4	-1045.7
AMSOL (SM5.4 water)	-827.0	-747.6	-797.5	-747.6	-827.0

Figure 5.22 - Heats of formation (PM3), in kJ mol^{-1}, relative to the symmetrical intermediate for 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Structure A)



Structure B	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
PM3					
MOPAC (No solvent)	-563.2	-545.0	-600.7	-545.0	-563.2
AMSOL (No solvent)	-561.7	-543.4	-599.4	-543.4	-561.7
MOPAC (COSMO)	-1045.7	-938.8	-1008.5	-938.8	-1045.7
AMSOL (SM5.4 water)	-827.0	-749.8	-810.9	-749.8	-827.0

Figure 5.23 - Heats of formation (PM3), in kJ mol^{-1}, relative to the symmetrical intermediate for 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Structure B)

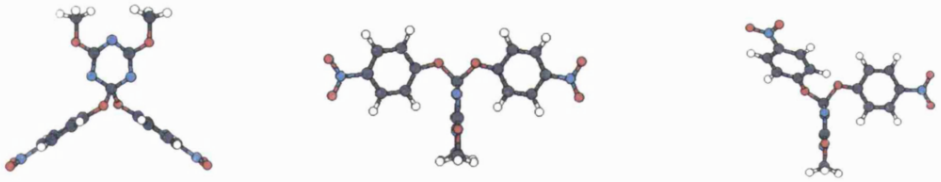


Structure C	Energy ΔH_f				
Computational method	Reactants	Transition State 1	Intermediate	Transition State 2	Products
PM3					
MOPAC (No solvent)	-563.2	-533.7	-607.1	-545.0	-563.2
AMSOL (No solvent)	-561.7	-543.4	-605.6	-543.6	-561.7
MOPAC (COSMO)	-1045.7	-932.4	-1002.2	-932.4	-1045.0
AMSOL (SM5.4 water)	-827.0	-737.1	-810.9	-749.9	-827.0

Figure 5.24 - Heats of formation (PM3), in kJ mol^{-1}, relative to the intermediate for 4-nitro-phenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Structure C)

'Intermediate' structures

A (C_2) B (C_{2v}) C (C_1)



Computational method	Intermediate Structure A	Intermediate Structure B	Intermediate Structure C
AM1	-265.4	**-250.7**	**-263.9**
PM3	**-604.8**	**-600.0**	-606.3
AMSOL (AM1/SM5.4A)	**-433.9**	**-462.3**	-466.1
AMSOL (PM3/SM5.4P)	**-797.6**	-810.9	-810.9
MOPAC (AM1/COSMO)	**-559.0**	**-591.2**	-594.5
MOPAC (PM3/COSMO)	**-992.0**	-1008.5	**-1005.8**

Figure 5.25 - Diagrams of the three possible structural isomers for the intermediate of the reaction of 4-nitrophenolate anions with 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine, as calculated by AMSOL (AM1). The energies given are the gas phase enthalpies of formation (averaged over the energies calculated in the two programs), solvated enthalpy of formation (COSMO) and solvated enthalpy of formation with Gibbs free energy of solvation (SM5.4A). The lowest energy structure as calculated by a particular method is shown in bold and all energies are shown in kJ mol^{-1} to one decimal place.

$C_6H_4(NO_2)O^- + C_3N_3(OMe)_2OPhNO_2 \rightarrow C_3N_3(OMe)_2(OPhNO_2)_2^- \rightarrow C_6H_4(NO_2)O^- + C_3N_3(OMe)_2OPhNO_2$

Reactant/Product structures 'Intermediate' structures

A (C_2) B (C_{2v}) C (C_1)

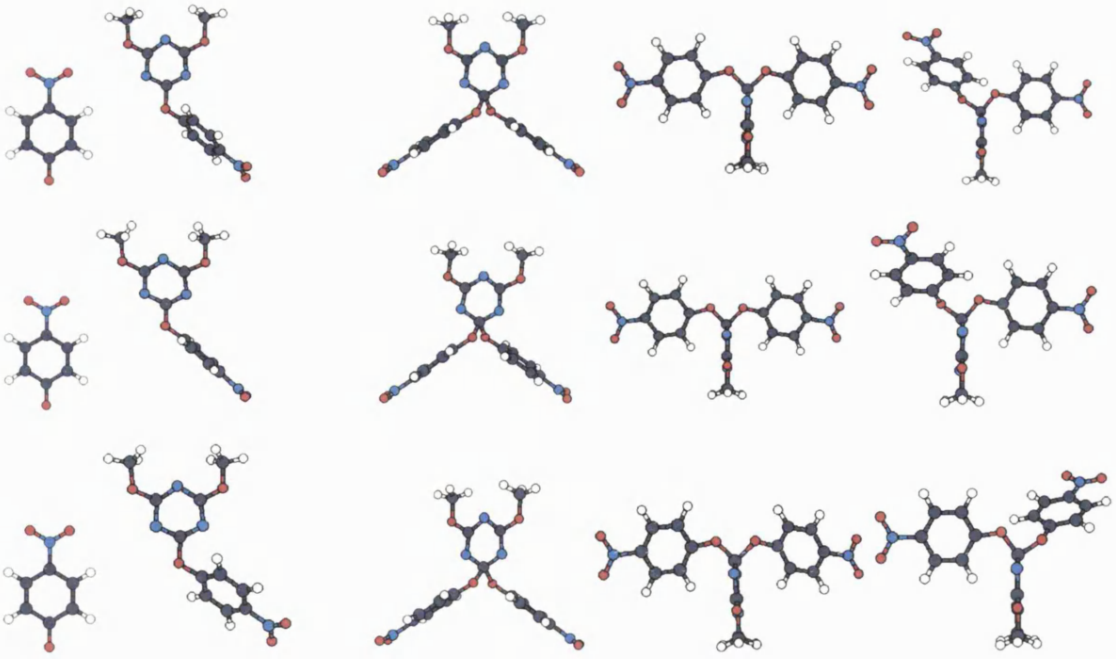


Figure 5.26 – Diagrams of the reactant, product and intermediate structures for the 4-nitrophenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine using the AM1 (top), AM1/COSMO (middle) and AM1/SM5.4A (bottom) methods.

5.3) Overview and Conclusions

The three Reactions (1, 2 and 3) previously studied experimentally have now been investigated computationally. Comparison of the calculated results with experiment allows evaluation of the various available semiempirical methods, with and without a continuum to represent the effects of solvation. The methods used were gas phase AM1 and PM3 and aqueous AM1/COSMO, PM3/COSMO, AM1/SM5.4A, PM3/SM5.4P calculations using the MOPAC93 and AMSOL6.1 programs.

It is noteworthy that the lowest energy structure was not the most symmetrical structure for all these reactions. However, the reactions studied all predicted similar mechanisms, albeit with different energetics. The calculation of the energy of all the most probable intermediate structures for a reaction is therefore desirable, in order to find the lowest energy reaction path.

All the reactions were computationally predicted to be stepwise. However one of the reactions was found experimentally to be concerted. The 3-nitrophenolysis of 3'-nitrophenyl-2,4,6-trinitrophenyl ether (Reaction 1) was examined by the two aqueous AM1 methods, AM1/SM5.4A and AM1/COSMO, only. Three conformers were calculated, these found a stepwise mechanism for both the methods as expected for a trinitrophenyl reaction, with a deep energy well (greater than 100 kJ mol^{-1}) indicating a long lived intermediate. The lowest energy conformer and therefore the lowest energy reaction path differed depending on the method used. The pyridinolysis of the 1-(4,6-diphenoxy-1,3,5-triazin-2-yl) pyridinium cation (Reaction 2) was examined by all the available computational methods; that is to say AM1, PM3, AM1/COSMO, PM3/COSMO, AM1/SM5.4A and PM3/SM5.4P. These show a stepwise mechanism with a small energy well (less than 10 kJ mol^{-1}) for the aqueous AM1 and a significant (22-33 kJ mol^{-1}) energy well for the aqueous PM3 calculations; this agrees with the experimentally predicted mechanism. However, this is within the error quoted for the ionic species of the parameterisation set for the continuum approaches. One reason for the destabilisation of the intermediate was the close proximity of the two pyridinyl nitrogens. The 4-nitrophenolate anions with 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine (Reaction 3) was examined by the same methods as Reaction 2. These methods all predict a stepwise mechanism contrary to the experimentally

found concerted mechanism. The energy wells for the three conformers were 21-34 and 26-74 kJ mol^{-1} for the aqueous AM1 and PM3 calculations respectively. The lowest and therefore most significant energy well, structure C, was outside the errors quoted in the methods used.

The two programs produced similar although differing gas phase AM1 and PM3 energies because of the different optimisation procedures used. Both gaseous methods cannot be directly compared to experiment as the experiment was carried out in solution. There are no preferences that can be drawn between these methods based on the results presented here; however AM1 has been found to be favourable for similar organic compounds previously and was therefore used in Chapter 4. As neither AM1 nor PM3 was extensively shown as better than the other further work should use AM1, as it has been used previously and would give a more direct comparison to those results.

The aqueous phase methods showed more instability during optimisation than the gas phase. The optimisation procedure for AMSOL SM5.4*x* was far more robust than the MOPAC COSMO methods. This is reflected in the high gradient norms, errors in the optimisation, than the other methods and produced spurious minor imaginary frequencies during optimisation. COSMO takes less time to compute than SM5.4*x*; however, due to the time needed to eliminate (or attempt to eliminate) the spurious imaginary frequencies, SM5.4*x* was quicker overall. Therefore, AMSOL SM5.4*x* is preferable to MOPAC93 COSMO.

The inclusion of a continuum adds a sizable error especially for the ionic species, this error should however be substantially reduced over the full reaction profile as the systematic error is likely to cancel out as the transition and intermediate structures are similar in structure. For example, the difference between the intermediate and transition structure, which have the same overall charge, is the lengthening one bond by around 0.4 Å.

Reactions 1 and 2 extensively agreed with the published experimental results both predicted stepwise mechanisms. Reaction 3 is also computationally predicted to be stepwise, however this is contrary to the experimentally predicted concerted mechanism. The well depth for this reaction energy profile was outside the quoted

errors for the continuum approaches. There are several possible reasons for this discrepancy between computational and experimental prediction;

- the computational (semiempirical continuum) method is not sufficiently accurate.
- the experimental prediction is incorrect or flawed.
- the assumptions used in the computational method were not valid.

There is a discrepancy between the solvent used in the computational (where the solvent is considered as pure water) and experimental (where the solvent is a 1:9 dioxan-water mixture was used, for Reactions 2 and 3). The dioxan was included experimentally in order to help dissolve the solute and is, as such, probably significant in the local solvation environment for the reactions. As dioxan is an organic solvent with a lower permittivity than water, this would reduce the overall dielectric of the solvent. An explicit representation of the solvent would help to better represent this effect although explicit solvation methods do have problems as well. As dioxan is an both an organic solvent and larger than water it is likely to have the most effect on solvating the more open structures (with larger solvent accessible surface area) around the hydrophobic ring systems.

A further complication is the change in the distribution of charge along the reaction pathway. A small molecule with a high charge density will be disproportionately stabilised by the solvent compared to a large molecule with the same charge but with a lower charge density. This will not produce a large energy difference between the structurally similar intermediate and TS structures, although it would have a marked effect between the reactants and the TS and intermediate species.

Another possibility considered is that the triazine ring may not behave aromatically and not proceed by a classical nucleophilic aromatic substitution mechanism. This can be disregarded however, as the structures produced in all three reactions are consistent with a nucleophilic aromatic substitution mechanism. The computational method used is also not biased towards a specific mechanism.

Semiempirical continuum methods proved unable to reproduce all three experimentally derived mechanisms. This casts doubt on mechanisms studied using this approach and calls for further testing of the available computational methods. The ability of all available methods to reliably reproduce the energetics

of the species at reasonable computational cost needs to be assessed. This can be achieved, for example, by reproducing experimental deprotonation energies of similar molecules.

Chapter 6

Gas phase protonation of 2-, 3-, 4-mono substituted pyridines and phenols using semiempirical and density functional approaches

How good is the computational treatment of the gas phase energetics of the pyridine and phenolate nucleophiles?

6.1) Introduction

The accuracy of theoretical methods to describe a molecule is of the utmost importance to computational chemists; this can be established by comparing calculated results to experimental data. The study of the thermodynamics of proton transfer reactions is of considerable interest to chemists as a large number of organic compounds are potential proton acceptors/donors. Thus the reproduction of experimental thermodynamic data of proton transfer reactions is a chemically interesting way of evaluating computational techniques. The ability of semi-empirical, AM1[40] and PM3, [45]and density functional methods, the BLYP and B3LYP density functionals, [109, 110]to reproduce experimental proton affinities (PA) for a series of pyridines[111] and phenols[112] is assessed here.

The thermodynamics of proton transfer for pyridines and phenols have been widely studied experimentally, although there are fewer studies of substituted phenols/phenolates than of pyridines/pyridiniums due to experimental difficulties. [113]The abundance of experimental thermodynamic data and the comparatively small size of the compounds make these systems ideal to evaluate computational methods. Furthermore, it is convenient to study pyridines and phenols as they are similar to the nucleophiles and leaving groups for the dye systems studied in chapter 5. Accordingly proton transfer in phenols and pyridines have been studied previously using various computational methods, for example PAs have been calculated for pyridines and phenolates using semiempirical methods[113-115] and *ab initio* (STO-3G[116] through to G2(MP2,SVP)). [117] Both semiempirical and high level *ab initio* calculations have reproduced experimental PA well, with the G2(MP2/6-31G*) method shown to be accurate to within 10kJ mol^{-1}.

6.2) Theory

The proton affinity (PA) of a base B is the enthalpy of deprotonation of its conjugate acid BH^+ in the gas phase, reaction 1, and may be determined from the standard enthalpies of formation of B, BH^+, and the proton at 298.15K according to eq. 6.1.

$$(BH^+)_g \rightarrow (B)_g + (H^+)_g \quad (1)$$

$$PA(B) = \Delta H°_f(B) - \Delta H°_f(BH^+) + \Delta H°_f(H^+) \quad (6.1)$$

In order to obtain a PA by means of semiempirical and density functional calculations, it is usual to employ the experimental value[118] of 1536.3 kJ mol^{-1} for $\Delta H°_f(H^+)$ in eq. 6.1, together with calculated values for $\Delta H°_f(B)$ and $\Delta H°_f(BH^+)$. The practical expression used for these calculations is therefore given by eq. 6.2, where the calculations are performed here using either the AM1 or the PM3 parameterisation of the semiempirical Hamiltonian or Density Functional Theory (DFT). The reliability of these methods can be evaluated by comparing to the relevant experimental energies.

$$PA(B)_{calc} = [\Delta H°_f(B) - \Delta H°_f(BH^+)]_{calc} + 1536.3 \text{ kJ mol}^{-1} \quad (6.2)$$

The gas-phase basicity (GB) is the sum of the proton affinity and temperature–entropy term ($-T\Delta S°$) for the deprotonation $BH^+ \rightarrow B + H^+$ in the gas phase. The temperature-entropy term is calculated via statistical partition function methods. [119] The free energy is given here as a correction to the total energy.

$$GB(B) = PA(B) - T\Delta S° \quad (6.3)$$

6.3) Methodology

The energies presented are for geometry optimised structures of all compounds in which symmetry was used wherever appropriate. All the systems were specified in Z-matrix co-ordinates; dummy atoms were included where three atoms were in a line, as for the cyano-substituted systems.

The AM1 and PM3 gas phase enthalpies of formation for substituted pyridines and phenols were performed using MOPAC93[74] and AMSOL6.1; [53]these produced identical values therefore only MOPAC values have been quoted. In all the semiempirical calculations the keywords GNORM=0.0 (MOPAC), GCOMP=0.0

(AMSOL), KICK=3 (AMSOL), DDMIN=0.0, SCFCRT=1E-10 and HESS=3 were specified. This increases the convergence criteria substantially over the default. The eigenvector following optimiser[120-123] was used in MOPAC calculations wherever possible. PAs were calculated using semiempirical methods, for a total of 28 pyridines, Figure 6.1, and 30 phenolates, Figure 6.2, containing a single substituent X in either the 2, 3, or 4 position.

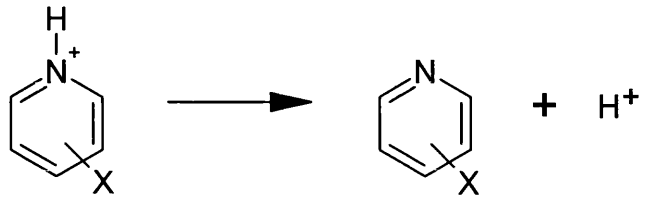


Figure 6.1 – The pyridine PA reaction examined for the decomposition of a substituted pyridinium cation to form a substituted pyridine and a proton.

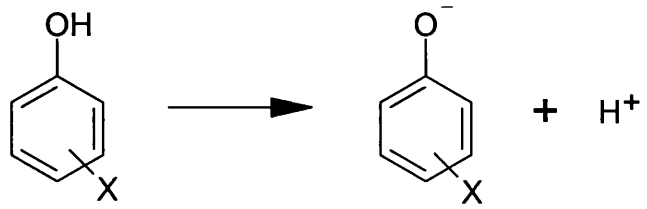


Figure 6.2 – The phenolate PA reaction examined for the decomposition of substituted phenol to form a substituted phenolate anion and a proton.

The density functional calculations were performed using the Gaussian 98 package. [46]The combinations of density functionals and basis sets used for all the systems were as follow: BLYP/6-31+G(d), BLYP/6-311++G(d,p), B3LYP/6-31+G(d) and B3LYP/6-311++G(d,p). Additionally the BLYP/6-31+G and BLYP/6-31G(d) methods were used for the phenolate systems. The (opt=Tight) restriction was initially used for all the systems, which significantly increases the convergence criteria for the optimisations; however it was removed for the small number of systems that failed to converge, namely the compounds with nitro-substituents. Seven phenols and seven pyridines with an even spread of experimental proton affinities were calculated using DFT, and the same set were used in calculations of gas phase basicities.

6.4) Results and discussion

6.4.1) Proton Affinity

Semiempirical computational studies of PAs have been published for substituted phenolates by Voets and co-workers[113, 115], and substituted pyridines by Szafran and co-workers. [114]The former was carried out on a different set of pyridines that additionally included the 2-COMe, 3-COMe and 4-COMe substituents and omitted the 4-Br and 4-CHO substituents. These studies consistently showed that AM1 reproduced experimental values better than PM3, with both correlation coefficients and gradients closer to unity and with a lower Root Mean Square (RMS) deviation. These published studies are plotted against experimental data with semiempirical and DFT calculated PA from this study. Discussion of overlapping results from previous studies will be only given where appropriate.

The calculated proton affinity (PA_{calc}) (eq. 3) is plotted against experimental proton affinity (PA_{expt}) for all of the methods. Semiempirical (AM1 and PM3) PA_{calc} is plotted against PA_{expt} for 29 phenolates and for 31 pyridines in Figure 6.4 and Figure 6.6 respectively. Semiempirical (AM1 and PM3) and DFT PA_{calc} is plotted versus PA_{expt} for 7 selected phenolates and 7 selected pyridines in Figure 6.5 and Figure 6.7 respectively. As the semiempirical PA is plotted alongside the DFT PA, a meaningful comparison can be drawn between the semiempirical and DFT calculations. The 7 phenolates and 7 pyridines chosen for study using DFT are a subset of those studied using semiempirical approaches with evenly distributed PA values. Regression analysis of these plots for all the semiempirical and DFT methods are summarised in Table 6.1. It was found that regression analysis is highly dependent on which substituents included and significantly changes occur if one is added or removed. As the values are spread over a larger range and have a high correlation coefficient the addition and removal of substituents has a lesser effect on the gas phase enthalpies of formation (ΔH_f), Section 6.4.3.

Table 6.1 – Regression analysis of experimental verses calculated Proton Affinities, in kJ mol^{-1}.

Figure Number	Symbol	Class of compound	Method	Number of compounds	Gradient	Intercept	Correlation coefficient (r)	Root Mean Square error
6.4	●	Phenols	AM1	29	0.988	6.7	0.925	16.2
6.4	■	Phenols	PM3	29	0.944	55.6	0.883	28.8
6.4	◆	Phenols	AM1 (Voets *et al.*)	28	0.991	4.4	0.939	13.8
6.5	●	Phenols	AM1	7	0.936	80.9	0.951	22.9
6.5	■	Phenols	PM3	7	0.891	126.0	0.744	14.4
6.5	●	Phenols	BLYP/631+G	7	1.200	-335.1	0.980	23.2
6.5	■	Phenols	BLYP/631G(d)	7	1.195	-260.9	0.997	47.8
6.5	·	Phenols	BLYP/631+G(d)	7	1.113	-204.4	0.991	15.7
6.5	◆	Phenols	BLYP/6311++G(d,p)	7	1.094	-165.2	0.996	5.3
6.5	+	Phenols	B3LYP/631+G(d)	7	1.104	-177.8	0.997	5.2
6.5	□	Phenols	B3LYP/6311++G(d,p)	7	1.081	-132.1	0.991	14.1
6.6	●	Pyridines	AM1	31	0.893	70.2	0.955	30.4
6.6	■	Pyridines	PM3	31	0.766	165.2	0.952	52.1
6.6	+	Pyridines	AM1 (Szafran *et al.*)	30	0.893	70.4	0.942	30.7
6.6	□	Pyridines	PM3 (Szafran *et al.*)	30	0.771	161.7	0.947	51.3
6.7	●	Pyridines	AM1	7	1.034	-65.7	0.992	34.2
6.7	■	Pyridines	PM3	7	0.864	71.5	0.983	57.6
6.7	·	Pyridines	BLYP/631+G(d)	7	1.084	-83.5	0.999	5.7
6.7	◆	Pyridines	BLYP/6311++G(d,p)	7	1.087	-78.6	0.999	5.0
6.7	+	Pyridines	B3LYP/631+G(d)	7	1.098	-93.6	0.999	4.3
6.7	□	Pyridines	B3LYP/6311++G(d,p)	7	1.102	-89.2	0.999	7.8

Table 6.2 – Regression analysis of experimental verses calculated Gas Basicities, in kJ mol^{-1}.

Figure Number	Symbol	Method	Number of compounds	Gradient	Intercept	Correlation coefficient (r)	Root Mean Square error
6.8	●	BLYP/631+G	7	1.225	-367.3	0.981	17.0
6.8	■	BLYP/631G(d)	7	1.221	-293.9	0.999	82.9
6.8	▲	BLYP/631+G(d)	7	1.136	-236.9	0.992	76.4
6.8	◆	BLYP/6311++G(d,p)	7	1.128	-210.0	0.998	63.8
6.8	+	B3LYP/631+G(d)	7	1.156	-245.4	0.987	61.5
6.8	□	B3LYP/6311++G(d,p)	7	1.102	-162.8	0.991	49.9
6.9	▲	BLYP/631+G(d)	6	1.057	-31.0	0.999	21.3
6.9	◆	BLYP/6311++G(d,p)	6	1.064	-29.8	0.999	29.4
6.9	+	B3LYP/631+G(d)	6	1.074	-43.9	0.999	24.0
6.9	□	B3LYP/6311++G(d,p)	6	1.083	-44.8	1.000	32.2

Table 6.3 – Regression analysis of the semiempirical gas phase enthalpies of formation of pyridines and pyridiniums, in kJ mol^{-1}.

Figure Number	Symbol	Class of compound	Method	Number of compounds	Gradient	Intercept	Correlation coefficient (r)	Root Mean Square
6.10	●	Pyridines	AM1	30	0.968	4.0	0.997	17.5
6.10	■	Pyridines	PM3	30	0.970	-11.5	0.998	19.3
6.11	+	Pyridinium ions	AM1	29	0.944	68.5	0.996	38.0
6.11	□	Pyridinium ions	PM3	29	0.953	68.2	0.998	40.5

The experimental PAs cover a range of over 125 kJ mol^{-1} for the phenolates and over 112 kJ mol^{-1} for the pyridines; however, these values are bunched up away from the origin. The intercept was highly gradient dependent that is to say that a small change in the gradient produces a large change in the intercept. Therefore, the intercept shall not be discussed here. Throughout the semiempirical calculations the total energy for each molecule was lower for most of the pyridines, pyridinium ions, phenol, phenolate ions than those in previous studies published. [113-115, 124, 125]

The plot of the semiempirical calculated PA of phenolate ions, Figure 6.4, shows that the AM1 parameterisation of the semiempirical Hamiltonian is significantly closer to experiment than PM3, for these compounds. AM1 has a lower RMS error and both correlation coefficient and gradient closer to unity than PM3. Both the published and calculated AM1 PA showed good agreement to experiment with similar gradients, although Voets and co-workers did not include 4-hydroxybenzaldehyde. The PA of 2-carboxyphenolate shows a significant energy difference to the value quoted by Voet and co-workers; this structure is strongly affected by the strength of the intramolecular hydrogen bond, which stabilises both the phenol and phenolate. The phenolate and phenol are both lower in energy here. This is in part due to tight convergence criteria and as the neutral phenol was optimised using a C_s structure, as shown in Figure 6.3:

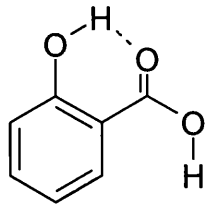


Figure 6.3 – The structure of 2-carboxyphenol calculated here.

The DFT calculated PA for the substituted phenolates, Figure 6.5, has a systematically higher sensitivity to the effect of the substituent compared with experiment, shown by the higher gradient, whereas semiempirically calculated PAs were systematically lower. AM1 had the nearest gradients to unity followed closely by B3LYP/6-311++G(d,p), BLYP/6-311++G(d,p), B3LYP/6-31+G(d), BLYP/6-31+G(d) and then PM3. The RMS error was lowest for B3LYP/6-31+G(d) and BLYP/6-311++G(d,p) at 5.2 and 5.3 kJ mol^{-1} respectively followed by B3LYP/6-311++G(d,p), PM3 and BLYP/6-31+G(d). The BLYP/6-31G(d) method

had the largest RMS error, indicating the importance of the diffuse function(s) used in the other DFT approaches, especially when an anion (phenolate) is calculated. The DFT calculations consistently had a much closer correlation coefficient that was much closer to unity than the semiempirical methods. In all PA calculations for the phenolates, AM1 outperformed PM3 at producing a good gradient and correlation coefficient, but these were both eclipsed by the DFT methods with larger basis sets, 6-31+G(d) and 6-311++G(d,p); the DFT methods also had excellent correlation coefficients.

AM1 showed a better agreement with experiment than PM3 for subsitituted pyridines, Figure 6.6. This was also shown by Szafran and co-workers published calculations[125]. The difference between Szafran and the calculations presented here is likely to originate from additional calculation of 4-bromopyridine and the high convergence criteria used here. The semiempirical methods underestimated the sensitivity to the effect of the substituent for the set of 31 pyridines chosen that is to say they had a gradient of less than one. AM1 reproduced experimental PA significantly better than PM3 for this group of compounds. Focusing on 7 pyridines, Figure 6.7, changes the correspondence between semiempirical methods and experiment markedly; AM1 has a near unity gradient and PM3 is closer but still gives the worst correlation of all of the approaches used. DFT, like AM1, overestimates the effect of the substituents consistently having a gradient greater than one for this set of pyridines with gradients ranging from 1.084 to 1.102. DFT has substantially lower RMS error than the semiempirical calculations of between 4.2 and 7.8 kJ mol^{-1} whereas AM1 and PM3 had RMS errors of 34.2 and 57.6 kJ mol^{-1} respectively. DFT also produced superb correlation coefficients, with all correlation coefficients above 0.999. Semiempirical methods also had a good correlation of 0.992 and 0.983 for AM1 and PM3 respectively. These observations are similar to those shown for the phenolates. Overall AM1 was shown to be the best semiempirical method for both phenols and pyridines. DFT produced markedly better results in terms of both correlation coefficients and root mean square errors than these computationally cheaper approaches. No DFT method with both diffuse and polarisation functions was shown to be significantly better than the rest, although the BLYP/6-311++G(d,p) was best overall. As BLYP/6-31+G(d) takes less computational time to calculate it provides a more computationally efficient result. One notable advantage of the DFT approaches,

as illustrated by the high correlation coefficient and low RMS errors, is the ability to produce consistent results for the all the phenolates and pyridines examined.

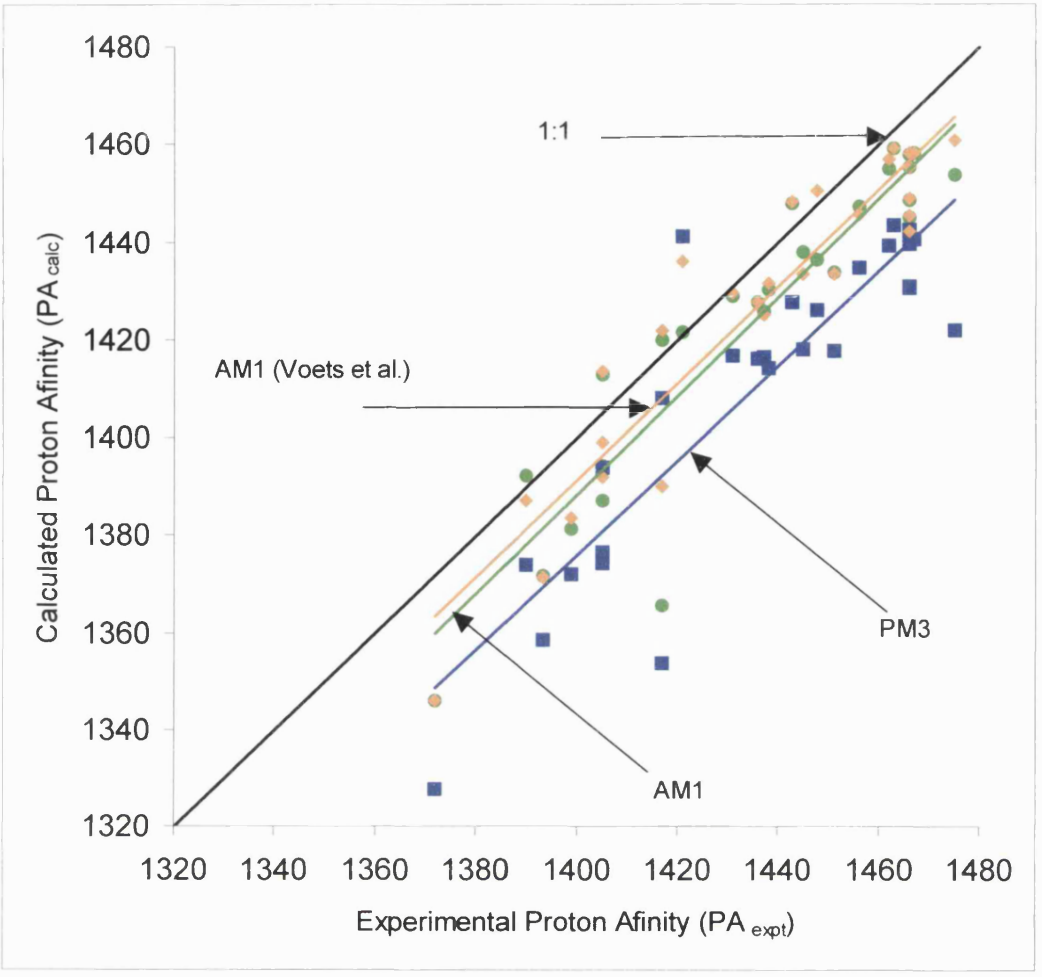


Figure 6.4 – Calculated AM1, PM3, Voets and co-workers[115] PAs for 29 substituted phenolate ions versus experimental PA[126-129] in kJ mol^{-1}.

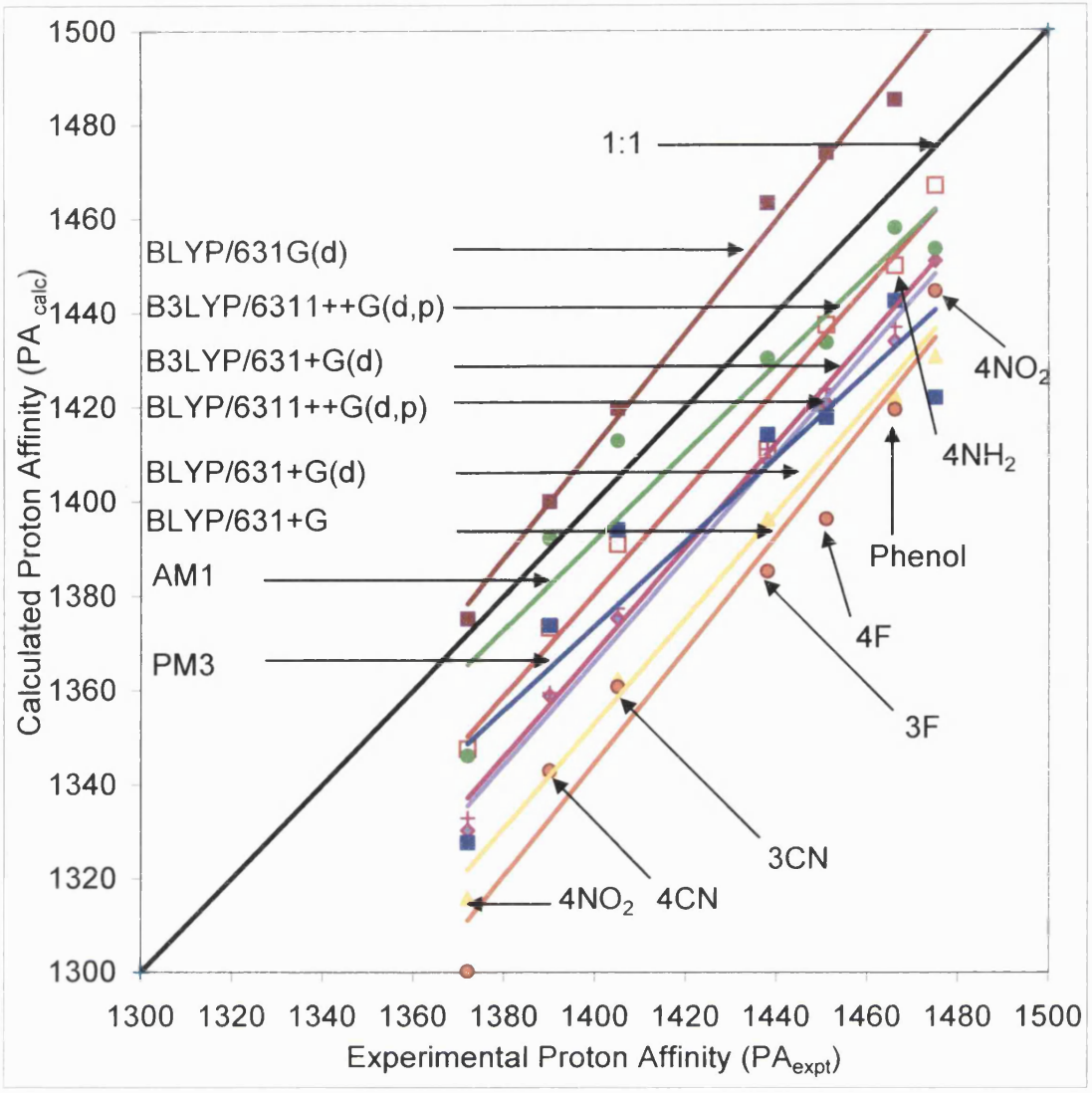


Figure 6.5 – Calculated AM1, PM3, BLYP and B3LYP PAs for 7 substituted phenolate ions versus experimental PA[127-129] in kJ mol^{-1}.

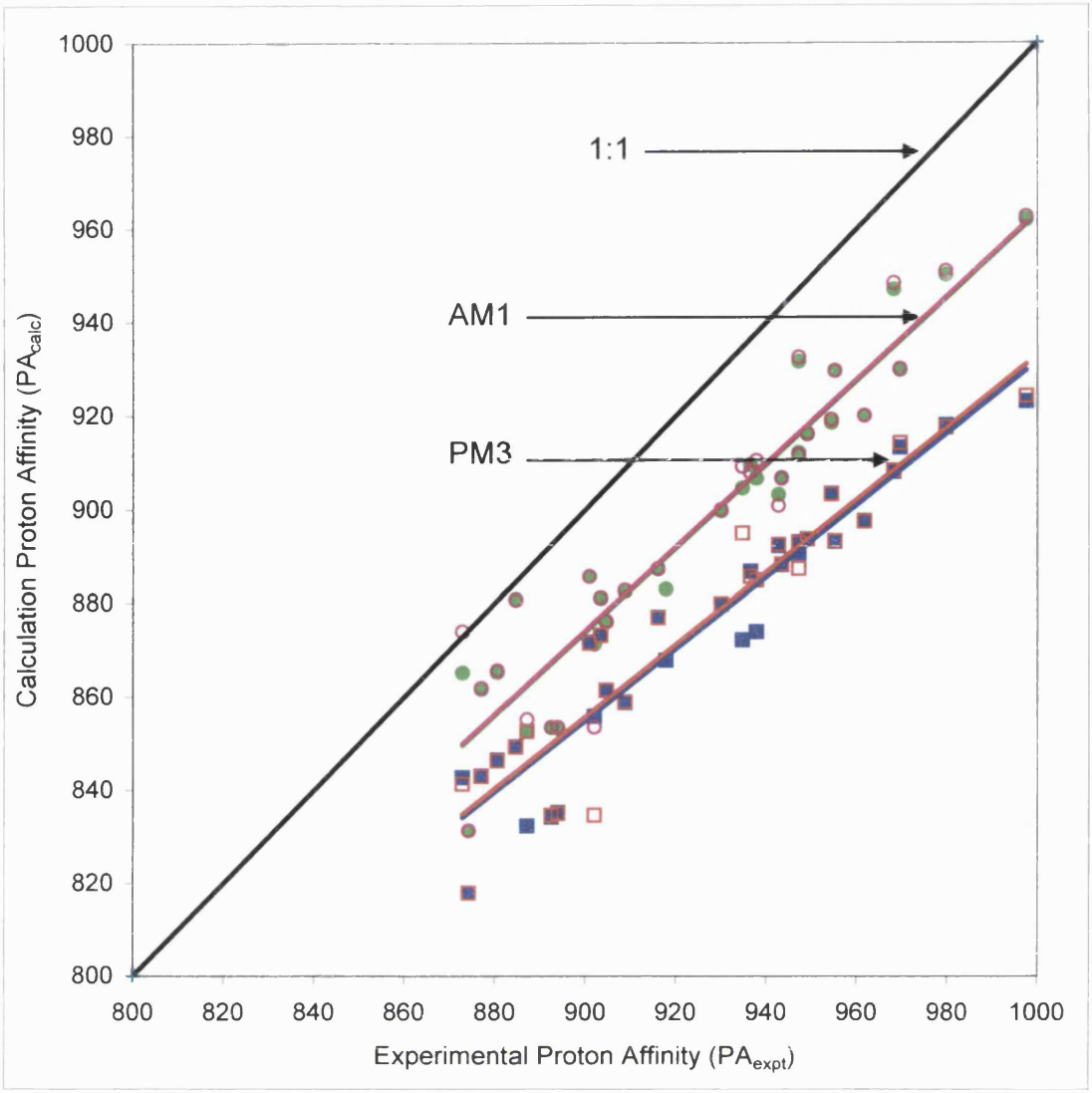


Figure 6.6 – Calculated AM1 and PM3 PAs for 31 substituted pyridines versus experimental PA[111, 130] in kJ mol^{-1}.

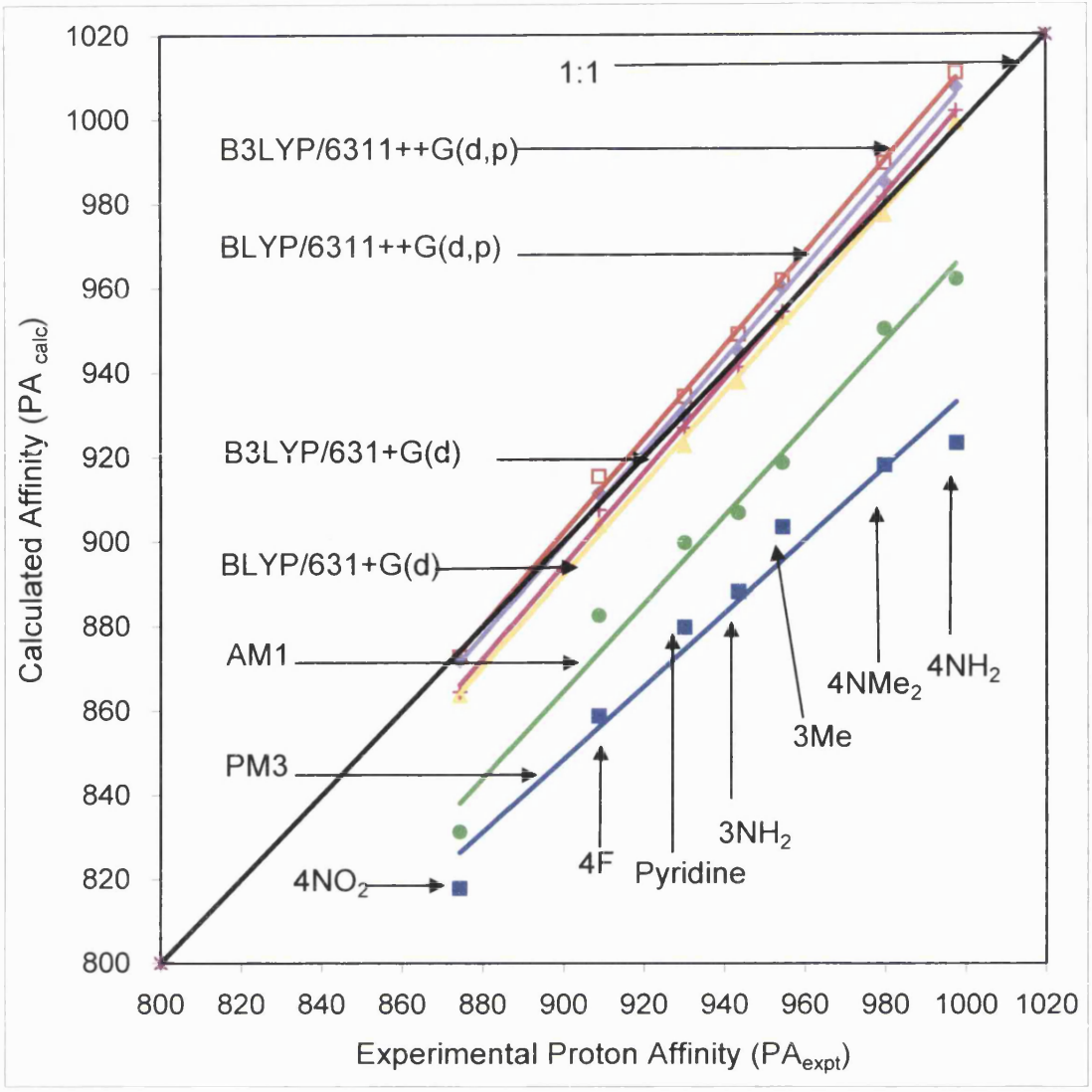


Figure 6.7 - Calculated AM1, PM3, BLYP and B3LYP PAs for 7 substituted pyridines versus experimental PA[111, 130] in kJ mol^{-1}

6.4.2) Gas Basicity

Plots of DFT calculated gas phase basicity (GB_{calc}) versus experimental (GB_{expt}) for the 7 selected phenolates and 6 selected pyridines are shown in Figures 6.8 and 6.9 respectively and regression analysis is summarised in Table 6.2. The regression analysis all the DFT methods show an oversensitivity to the effect of the substituents, with gradients above one. These gradients from the methods are organised into two groups; the methods with small basis sets BLYP/6-31+G and BLYP/6-31G(d), which have gradients of 1.225 and 1.221 respectively, and the rest that have gradients, which range from 1.102 to 1.156. The RMS error was large for most of the methods, due to a near consistent; with the exception of BLYP/6-31+G with a RMS of 17.0 kJ mol^{-1}, underestimation of the absolute GB compared to experiment. This underestimation caused the RMS error to become as high as 82.9 kJ mol^{-1} for BLYP/6-31G(d). A near unity, correlation coefficient was given for all the methods ranging from 0.981 for BLYP/6-31+G to 0.999 for BLYP/6-31G(d). The plot of the calculated (GB_{calc}) versus experimental (GB_{expt}) Gas phase Basicities for a set of six substituted pyridines using DFT methods, Figure 6.10 and Table 6.2, consistently have gradients above one. This, although less than for the phenolates, shows an oversensitivity to the effect of the substituent with gradients ranging from 1.057 to 1.083. The RMS errors range from 21.3 to 24.0 kJ mol^{-1} due to a systematic overestimation of the GB by the DFT methods. The correlation coefficient of 0.999 and 1.000 to three decimal places for all the DFT methods is outstanding and shows excellent consistency for these density functional approaches, like in the phenolates. Throughout both the pyridines and phenols, the effect of the substituent is overestimated, but not by a large amount. Overall, DFT provides a good estimation of gas basicity, with the best description given is when a balanced basis set with both diffuse and polarised functions are employed.

Most methods including density functional, Hartree-Fock and semiempirical approaches tend to overestimate vibrational frequencies. Empirically derived scaling factors are generally applied to correct this error. Vibrational frequency scaling factors have been calculated for various combinations of method/basis set. These are particularly significant in Hartree-Fock approaches but are less significant with density functional approaches. The scaling factors suitable for fundamental vibrations derived from a least-square fit of frequencies for

BLYP/631G(d) and B3LYP/631G(d) are 0.9945 and 0.9614 respectively. [131] The methods examined here include an additional diffuse function, however applying the larger error to the similar B3LYP/6-31+G(d) method does give an approximation of the effect of scaling factors on the overall energy. There are different components of the energy notably the Zero-Point Energy (ZPE), thermal energy and vibrational entropy affected by the vibrational frequencies. ZPE is much larger than the thermal correction which is in turn much greater than the Vibrational entropy.

In order to show the magnitude of the error the Gas Basicity of Phenolate using B3LYP/6-31+G(d) is examined. As ZPE is a simple linear function of frequencies the effect can be calculated: -

Effect on ZPE = Difference in ZPE correction of Phenolate and Phenol multiplied by scaling factor
= 36.8 $kJmol^{-1}$ x 0.96 (Scaling factor) = 35.4 $kJmol^{-1}$

The effect of the scaling factors on the ZPE being 1.4 $kJmol^{-1}$. The effect on the smaller thermal energy is non linear. Based upon typical values of vibrational frequencies for these type of compounds the percentage error in the thermal energy was found to be lower than the scale factor error. The effect is larger on the higher frequencies and has only a slight effect at lower frequencies. Due to the small magnitude of the thermal energy even if it was assumed to be linear the effect on the thermal energy for the gas basicity of Phenolate would be 0.03$kJmol^{-1}$. The remaining source of error, the vibrational entropy, is considerably smaller than the thermal energy term and therefore is not significant. Vibrational entropy is only significant when there are lots of very low frequencies this is not the case for the compounds studied here.

This effect is likely to be consistent for all the calculations using a particular method. An effect of this small magnitude is unlikely to affect the comparison of the graphs dramatically. Therefore, it is a satisfactory approximation to omit the effect of the scaling factors from this comparison.

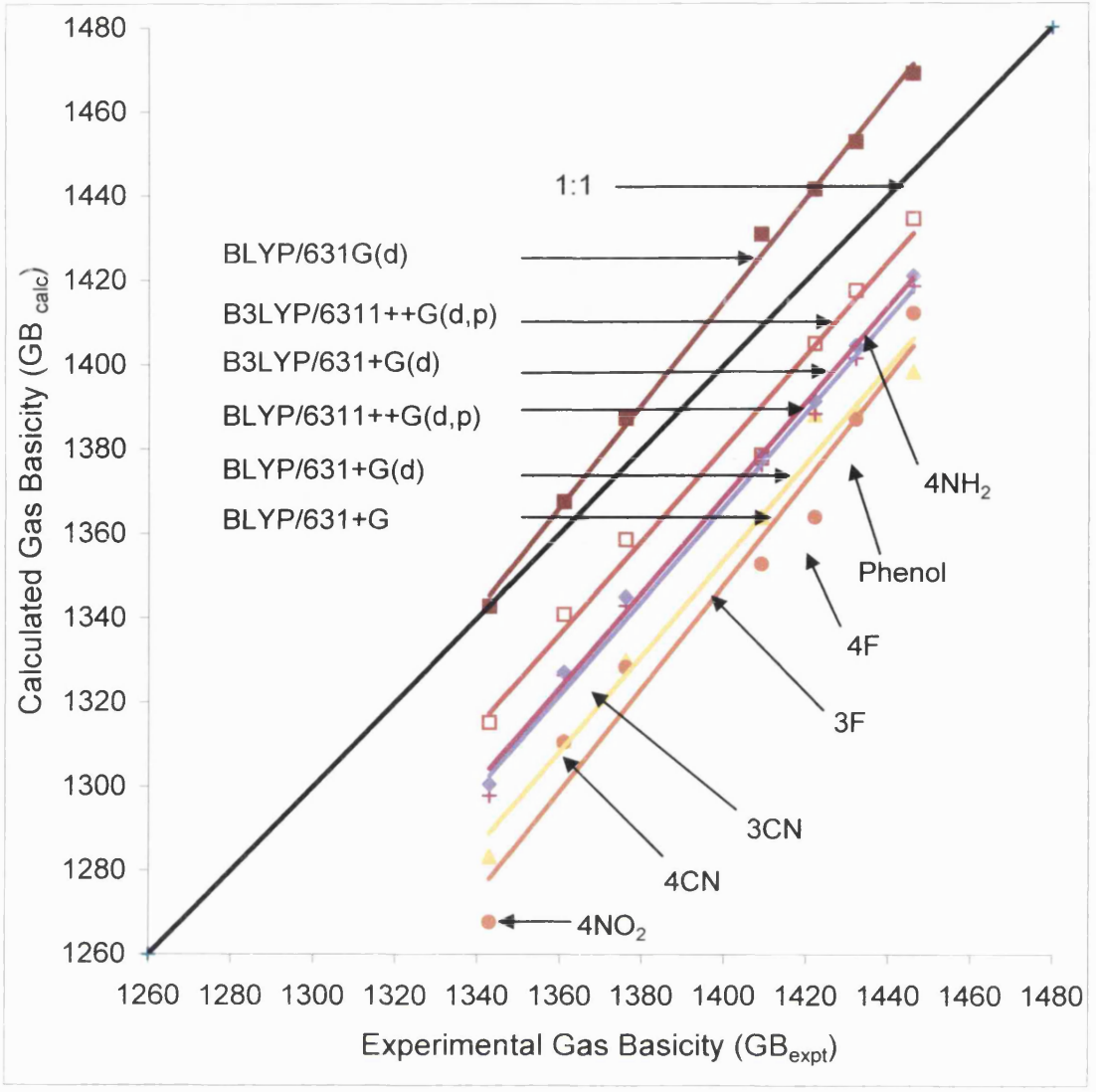


Figure 6.8 – Calculated BLYP and B3LYP GBs for 7 substituted phenolate ions versus experimental GB in kJ mol^{-1}.

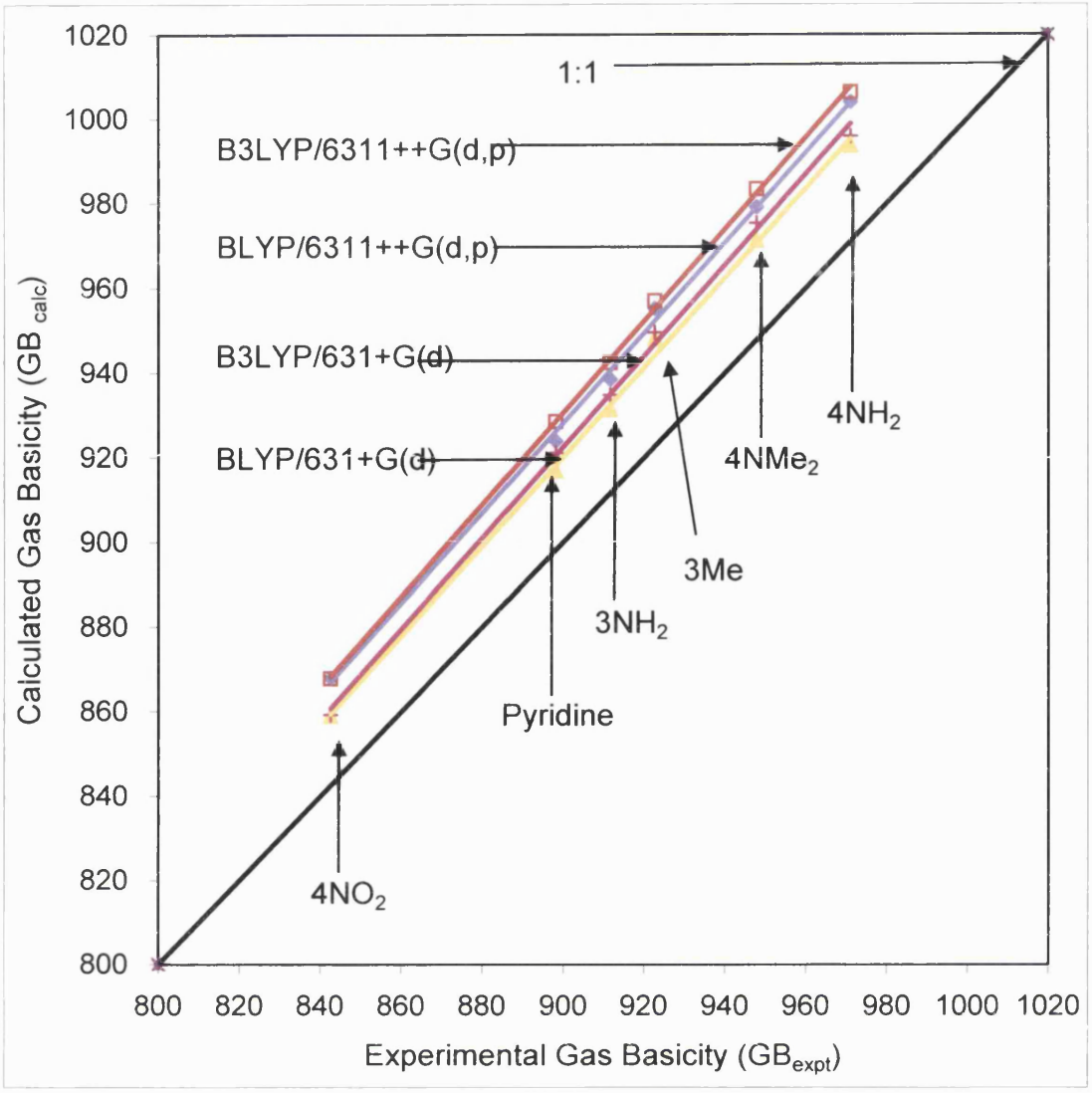


Figure 6.9 – Calculated BLYP and B3LYP GBs for 6 substituted pyrdines versus experimental GB in kJ mol^{-1}.

6.4.3) Enthalpies of formation of pyridine and pyridinium ions

Plots of the calculated versus experimental gas phase enthalpies of formation (ΔH_f) of substituted pyridine and pyridinium ions, the two calculated components of the PA, are shown in Figure 6.10 and Figure 6.11 respectively and regression analysis in shown in Table 6.3. The calculated enthalpies of formation ($\Delta H_{f,calc}$) of the individual pyridine and pyridinium cations, used in the calculation of the PA, were compared to experiment ($\Delta H_{f,expt}$). The semiempirical PAs of pyridines had higher RMS errors than the phenolates; also these had readily available enthalpies of formation. Analysing the components used to calculate the PA, the origin of the error could be assigned. The calculated enthalpies of formation of pyridines show an excellent agreement and pyridinium ions show a good correlation, Table 6.3, to experiment. This is shown, in Figures 6.10 and 6.11, where both the gradients and correlation coefficient are close to unity. The RMS error is relatively large, approximately 20 and 40 kJ mol^{-1} for the pyridines and pyridinium ions respectively. The RMS error in the PA therefore mainly originates from the calculation of the pyridinium ions, with calculated energies being systematically higher than experiment. This may be in part due to fewer ions being used in the parameterisation set for the semiempirical methods.

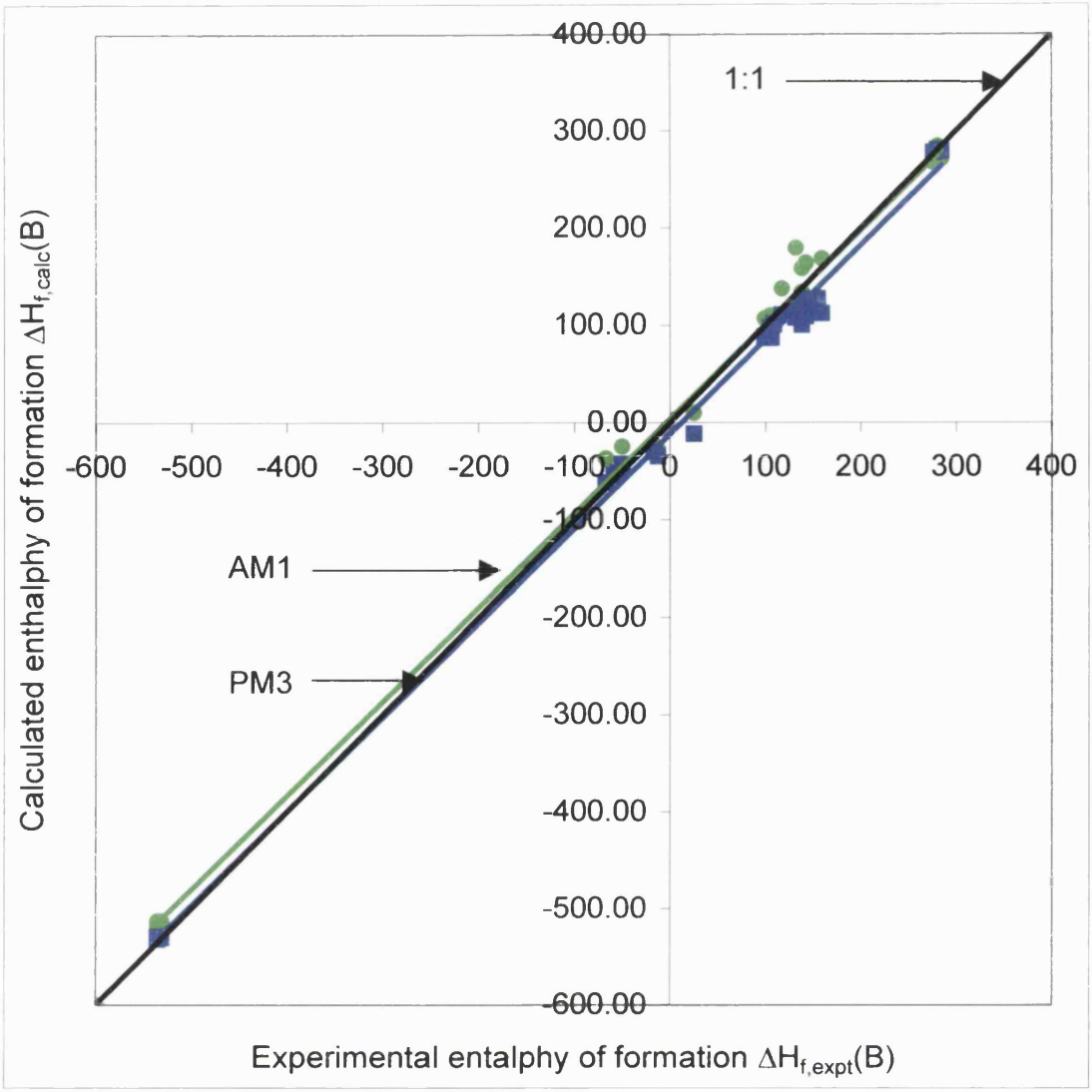


Figure 6.10 – Calculated semiempirical enthalpy for formation of 30 substituted pyridines versus experimental[114] in kJ mol^{-1}.

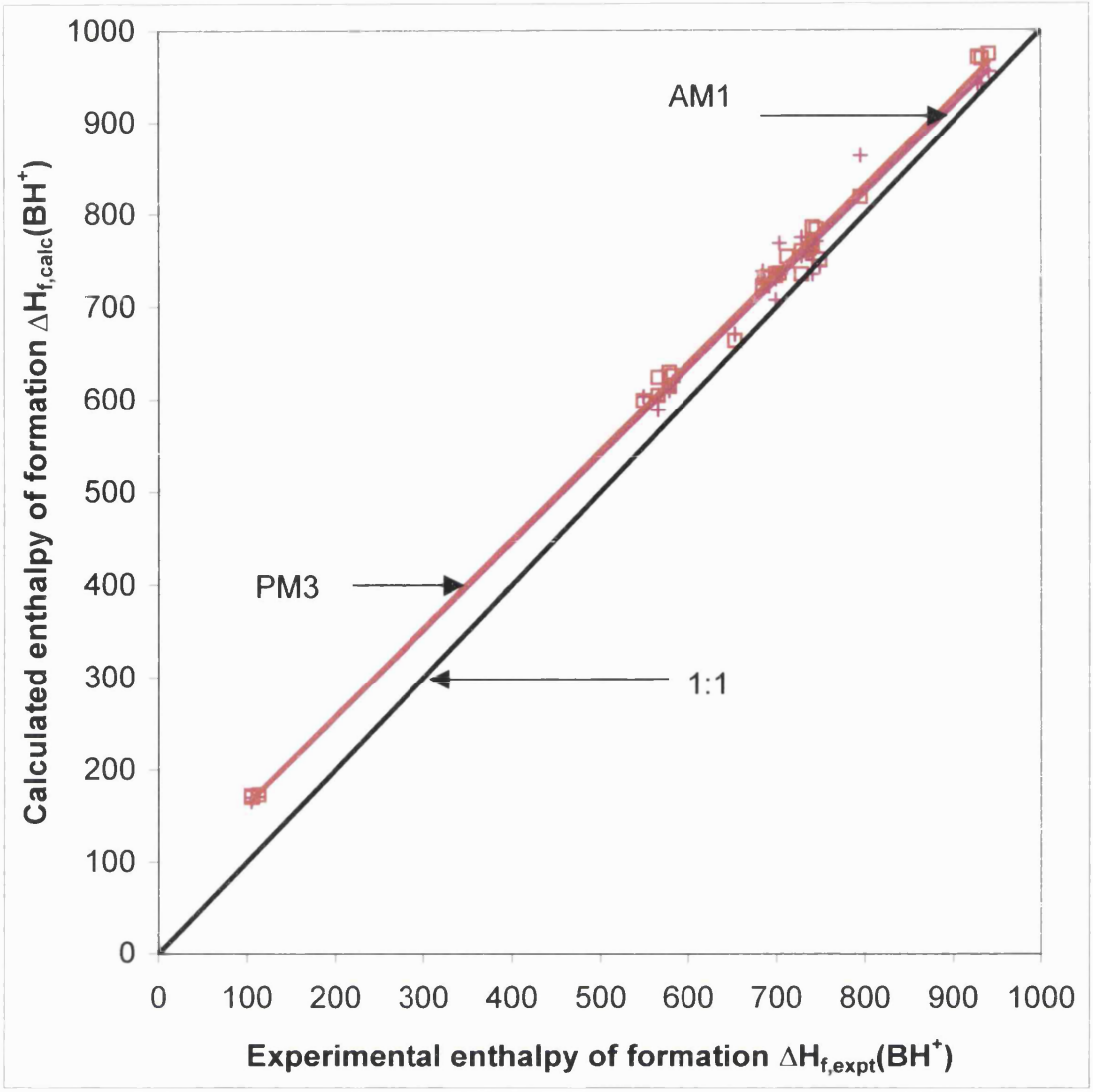


Figure 6.11 – Calculated semiempirical enthalpy for formation of 30 substituted pyridinium ions versus experimental[114] in kJ mol^{-1}.

6.5) Conclusion

AM1 reproduces experimental proton affinities better than PM3 for pyridines and phenolates. There is significant error in each method examined, which indicates that for the semiempirical methods, the majority of the error is within the calculation of the ionic species. Semiempirical methods tended to be less sensitive to the effect of the substituent and have a gradient against experiment of lower than one; conversely density functional methods overestimated this effect.

When plotted against experimental proton affinities, density functional calculated proton affinities generally produce gradients and correlation coefficients closer to unity with lower RMS errors than semiempirical methods. The semiempirical equivalent AM1 calculation does give comparable estimates although inconsistent results at considerably less computational cost. All the density functional methods used gave comparable results given a reasonable size basis set with both diffuse and polarised functions. Therefore the least expensive method with both diffuse and polarised functions BLYP/6-31+G(d) should efficiently give an adequate description of the gas phase energetics of similar organic compounds to pyridines and phenols.

6.6) Evaluation

Semiempirical methods and DFT reproduce gas phase energetics or pyridines and phenolates well. DFT gives a much superior gradient and correlation coefficient versus experiment and is therefore more consistent than the semiempirical approaches. As the majority of the calculations involved in a mechanism are similar in nature this consistency is crucial.

Chapter 7

Modelling aqueous dissociation energies of 2-, 3-, 4-mono substituted pyridines and phenols using semiempirical and density functional approaches

How good is the computational treatment of the solvated energetics of the nucleophiles?

7.1) Introduction

The accuracy of theoretical methods for describing a chemical reaction in solution is of the utmost importance to computational chemists. The accuracy of solvation models describing them is of considerable interest, as most chemical reactions occur in solution. As most organic compounds are potential proton acceptors/donators, the thermodynamics of proton transfer reactions are of major interest to chemists. Thus the reproduction of experimental thermodynamic data is an interesting means of evaluating theoretical techniques.

The calculation of the aqueous deprotonation energies requires the use of an accurate solvation model. Solvation models cover a large area of research with lots of texts [29, 47-49, 132]describing them in detail, therefore only a short summary has been given in this thesis, in Section 2.2. There are a wide variety of solvation models available; these vary greatly in accuracy and computational expense, for example COSMO, [52]SM5.4*x*, [107]GCOSMO, [133]COSMO-RS, [134, 135]PCM, [57, 58]SCIPCM, [136]SCRF, [137, 138]CPCM, [59]or explicit solvation models. [139]

The ability of the underlying solute descriptions, detailed in Section 2.2, (AM1, [40]PM3[45] and BLYP and B3LYP density functionals[109, 110]) to reproduce experimental proton affinity (PA) for a series of pyridines and phenols is shown in Chapter 6. The ability of semiempirical and Density Functional (DF) continuum computational approaches to reproduce pK_a and Gibbs free energy of dissociation for a pyridines and phenols is evaluated here. The approaches examined are the semiempirical MOPAC93[74]/COSMO[52] (conductor-like screening solvation

model) and AMSOL6.1[53]/SM5.4[107] (Solvation Model 5.4) and DF Gaussian 98[46]/CPCM[59] (COSMO Polarised Conductor Model). The COSMO and SM5.4 continuum models are based on the generalised Born approximation, as described in Section 2.2, in which the solute is enclosed in a cavity that is surrounded by a solvent. The solvent is modelled as an infinite polarisable continuum of a certain permittivity; the value used in these calculations was that equivalent to water. First solvation shell effects are included in the case of SM5.4*x*[107] perturbing the continuum based on the solvent accessible surface area. CPCM[59] models the solvent as an approximation to the Poisson equation that is solved for a grid of points on the solutes surface.

There are a limited number of published studies of aqueous proton transfer for pyridines and phenols using continuum solvation models. The effect of fluorine substitution of phenols has been examined by Urban and co-workers[124] using AM1 and PM3 combined with the SM2, [65]SM3[66] and GB/SA[140] continuum models. This study showed a good agreement between calculated and experimental Gibbs free energies of hydration; however only a small range of similar compounds was studied. The pK_a for carboxylic acids and phenols have been calculated using AM1/COSMO and AM1/SM2.1 by Schüürmann and co-workers; [141]these showed a reasonable correlation to experimental data for the compounds selected. The reproduction of pK_a of phenols using DFT/COSMO[142] and carboxylic acids using *ab initio*/PCM[143] concluded the methods were not sufficiently accurate to predict absolute pK_a values. The pK_a of substituted pyridines were studied using AM1/SM2, HF/6-31G(d)/IPCM, HF/6-31G(d)/IPCM, HF/6-31G(d)/SCIPCM and MP2/6-31G(d)/IPCM; these methods gave RMS errors compared with experiment of between 1.2 and 4.1 a pK_a units[144] (1 pK_a unit ≡ 5.71 kJ mol^{-1} in the standard Gibbs free energy). These methods are clearly not sufficiently accurate to predict absolute pK_a values. The main source of error cited in these papers is that of the solvation of the charged species, in particular the size of the cavity used which is related to the van der Waals radii of the atoms involved. [145]The use of very high level and computationally expensive *ab initio* complete basis set (CBS) and Gaussian-n methods combined with the CPCM solvation model has been shown to reproduce pK_a values of six carboxylic acids to within an accuracy of half a pK_a unit[146] (~2.8 kJ mol^{-1}). These levels of theory are not, at present, computationally feasible for general use. Density functional (DF) calculations yield more reliable molecular potentials than semiempirical methods,

[134]this suggests that the use of a DFs with a continuum method should be more accurate than semiempirical implementation of continuum models. These DF methods are less computationally intensive than *ab intio* CSB and Gaussian-*n* approaches but significantly more expensive to calculate than semiempirical dielectric continuum methods. It is important to establish what the relative accuracies of the approaches are in order to allow an informed choice of a computationally efficient method. The recent substantial increases in computer power and refinement of theoretical methods have allowed reasonably large solvated systems to be studied by DF methods.

7.2) Theory

The proton affinity PA of a base B is the enthalpy of deprotonation of its conjugate acid BH^+ in the gas phase, reaction 1, and may be determined from the standard enthalpies of formation of B, BH^+, and the proton at 298.15K according to eq. 7.1.

$$(BH^+)_g \rightarrow (B)_g + (H^+)_g \quad (1)$$

$$PA(B) = \Delta H°_f(B) - \Delta H°_f(BH^+) + \Delta H°_f(H^+) \quad (7.1)$$

It is usual to employ the experimental value of 1536.3 kJ mol^{-1} [118]for $\Delta H°_f(H^+)$ in eq.7.1, together with calculated values for $\Delta H°_f(B)$ and $\Delta H°_f(BH^+)$, as this is badly estimated using computational approaches. Practically the expression used is given by eq. 7.2.

$$PA(B)_{calc} = [\Delta H°_f(B) - \Delta H°_f(BH^+)]_{calc} + 1536.3 \text{ kJ mol}^{-1} \quad (7.2)$$

The gas-phase basicity (GB) is the sum of the proton affinity and temperature–entropy term ($-T\Delta S°$) for the ionization $BH^+ \rightarrow B + H^+$ in the gas phase, shown in eq. 7.3. The free energy is given here as a correction to the total energy. The temperature-entropy term is calculated via statistical partition function methods.

$$GB(B)_{calc} = PA(B) - T\Delta S \quad (7.3)$$

$PA(B)_{calc}$ and $GB(B)_{calc}$ are shown in chapter 6 for a set of 30 phenolates and 28 pyridines using semiempirical and 7 phenolates and 7 pyridines using Density Functional (DF) approaches .

The Gibbs free energy cycle for pyridine and phenolate anion studied here are shown in Figures 7.1 and 7.2 respectively.

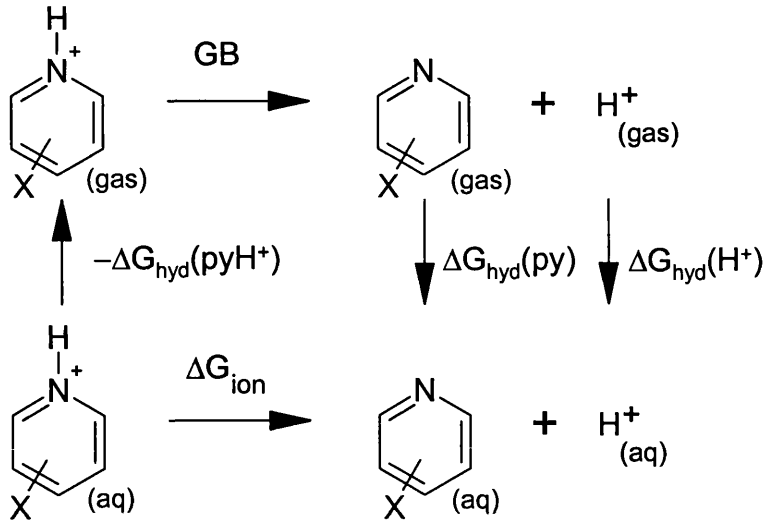


Figure 7.1 – The Gibbs free energy cycle for deprotonation of a *x*-substituted pyridinium cation to form a *x*-substituted pyridine and a proton.

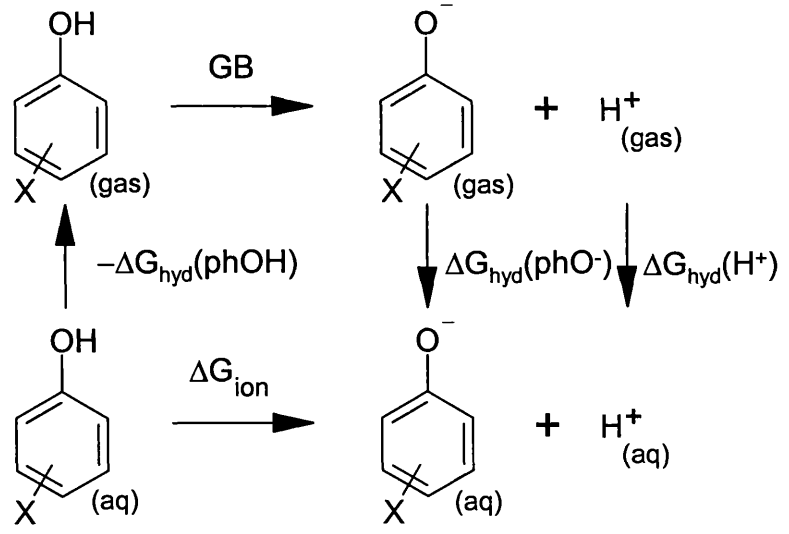


Figure 7.2 – The Gibbs free energy cycle for deprotonation of a *x*-substituted phenol to form a *x*-substituted phenolate and a proton.

The Gibbs free energy of deprotonation of an acid BH^+ in aqueous solution, reaction 2, may be determined from the standard aqueous Gibbs free energies of formation of B, BH^+ and H^+ at 298.15K where the standard state in solution is 1 mol dm^{-3} (1 M). Likewise, the enthalpy of deprotonation of an acid BH^+ in aqueous solution can be determined from the standard aqueous enthalpy of formation of B, BH^+ and H^+ at 298.15K.

$$(BH^+)_{aq} \rightarrow (B)_{aq} + (H^+)_{aq} \qquad (2)$$

The Gibbs free energy of deprotonation, $\Delta G°_{ion}$, is also the sum (eq. 7.4) of the GB (eq. 7.3) and the difference $\Delta\Delta G°_{hydration}$ in the Gibbs free energies of hydration of B, BH^+ and H^+, eq. 7.5.

$$\Delta G°_{ion} = GB + \Delta\Delta G°_{hydration} \qquad (7.4)$$

$$\Delta\Delta G°_{hydration} = \Delta G°_{hyd}(B)_{calc} - \Delta G°_{hyd}(BH^+)_{calc} + \Delta G°_{hyd}(H^+)_{expt} \qquad (7.5)$$

Analogously, the enthalpy of deprotonation, $\Delta H°_{ion}$, is also the sum (eq. 7.6) of the PA (eq. 7.2) and the difference $\Delta\Delta H°_{hydration}$ in the enthalpies of hydration of B, BH^+ and H^+, eq. 7.8.

$$\Delta H°_{ion} = PA + \Delta\Delta H°_{hydration} \qquad (7.6)$$

$$\Delta\Delta H°_{hydration} = \Delta H°_{hyd}(B)_{calc} - \Delta H°_{hyd}(BH^+)_{calc} + \Delta H°_{hyd}(H^+)_{expt} \qquad (7.7)$$

The Gibbs free energy of hydration $\Delta G°_{hyd}$ and enthalpy of hydration $\Delta H°_{hyd}$ for a substance A is given by eq. 7.8 and eq. 7.9 respectively.

$$\Delta G°_{hyd}(A) = \Delta G°_{aq}(A) - \Delta G°_{f}(A) \qquad (7.8)$$

$$\Delta H°_{hyd}(A) = \Delta H°_{aq}(A) - \Delta H°_{f}(A) \qquad (7.9)$$

It is usual to apply experimental values for the gas phase Gibbs free energy and gas phase enthalpy of formation and hydration of H^+, due to the bad estimations given by computational approaches.

These can be related to the aqueous Gibbs free energy (eq. 7.10) and aqueous enthalpy (eq. 7.13) of formation by rearranging eq. 7.8 and eq. 7.9.

$$\Delta G°_{aq}(H^+) = \Delta G°_f(H^+)_{expt} + \Delta G°_{hyd}(H^+) \quad (7.10)$$

where $\Delta G°_f(H^+)$ = 1517.0 kJ mol^{-1} (298.15, standard state 1 atmosphere) [118]and the value of $\Delta G°_{hyd}(H^+)$, the absolute Gibbs free energy of hydration for the proton, has recently been determined[147] as –1104.5 ± 0.3 kJ mol^{-1}.

$$\Delta H°_{aq}(H^+) = \Delta H°_f(H^+)_{expt} + \Delta H°_{hyd}(H^+) \quad (7.11)$$

where $\Delta H°_f(H^+)$ = 1536.3 kJ mol^{-1}[118] (as above) and the value of $\Delta H°_{hyd}(H^+)$, the absolute hydration enthalpy of the proton, has recently been determined[147] as –1150.1 ± 0.9 kJ mol^{-1}.

The parameters used for the solvation methods are for a single temperature of 298.15K. Therefore the values are all for that temperature and they cannot be performed at different temperatures, unlike in experimental studies. The solvation models are also parameterised to calculate specific thermodynamic quantities. The SM5.4A(water) and SM5.4P(water) solvation models[107] in AMSOL6.1[53], in combination respectively with AM1 and PM3, solute yield an approximation to the standard Gibbs free energy of formation $\Delta G°_{aq}$ in 1 Mol dm^{-3} aqueous solution at 298.15K. Therefore a practical expression used to calculate the Gibbs free energy of deprotonation of an acid in aqueous solution combining SM5.4*x* calculated aqueous Gibbs free energies of $\Delta G°_{aq}(BH^+)$ and $\Delta G°_{aq}(B)$ with the experimental aqueous Gibbs energy of formation for a proton from eq. 7.10, $\Delta G°_{aq}(H^+)$, is given by eq. 7.12.

$$\Delta G°_{ion} = [\Delta G°_{aq}(B) - \Delta G°_{aq}(BH^+)]_{calc} + 412.5 \text{ kJ mol}^{-1} \quad (7.12)$$

The CPCM[59] solvation model in Gaussian98[46] yields an approximation to the standard Gibbs free energy of hydration $\Delta G°_{hyd}$ in aqueous solution. A practical expression to calculate the Gibbs free energy of deprotonation of an acid in aqueous solution combining calculated GB and Gibbs free energy of hydration for B and BH^+ from CPCM and the Gibbs energy of hydration for a proton, $\Delta G°_{hyd}(H^+)$, is given by eq. 7.15. GB_{calc} is estimated from the free energy corrections to the total energy calculated from the gas phase vibrational frequencies as described in chapter 6.

$$\Delta G°_{ion} = GB_{calc} + -1104.5 \text{ kJ mol}^{-1} + \Delta G°_{hyd}(B)_{calc} - \Delta G°_{hyd}(BH^+)_{calc} \quad (7.13)$$

The conductor-like screening model (COSMO) [52]method implemented in MOPAC93[74] with a AM1 and PM3 solute description yields an estimate for the standard enthalpy of formation $\Delta H°_{aq}$ in aqueous solution. A practical expression used to calculate the enthalpy of deprotonation of an acid in aqueous solution combining COSMO calculated aqueous enthalpies of formation of $\Delta H°_{aq}(BH^+)$ and $\Delta H°_{aq}(B)$ with the enthalpy of formation for a proton from eq. 7.13, $\Delta H°_{aq}(H^+)$, is given by eq. 7.14.

$$\Delta H°_{ion} = [\Delta H°_{aq}(B) - \Delta H°_{aq}(BH^+)]_{calc} + 386.2 \text{ kJ mol}^{-1} \quad (7.14)$$

The experimental value of the Gibbs free energy of deprotonation of an acid in aqueous solution is obtained simply from the pK_a value of the acid by means of eq. 7.15 (for T = 298K).

$$\Delta G°_{ion} = 2.303RT\, pK_a = 5.71\, pK_a \text{ kJ mol}^{-1} \quad (7.15)$$

Gibbs free energies of hydration calculated by means of the SM5.4 and CPCM methods correspond to transfer of a solute from a 1 mol dm^{-3} ideal gas for the gaseous state to a 1 mol dm^{-3} ideal solution for the solution state, in accord with the recommendation of Ben-Naim. [148]Since most literature Gibbs free energies of hydration are for transfer of a solute from the gaseous state at 1 atmosphere to a 1 mol dm^{-3} solution, it is necessary to correct for the change of standard state before comparison can be made between SM5.4 calculated and experimental values of $\Delta G°_{hyd}$ at 298.15K, as given by eq. 7.16.

$$\Delta G°_{hyd}[1 \text{ M} \rightarrow 1 \text{ M}] = \Delta G°_{aq}[1 \text{ atm} \rightarrow 1 \text{ M}] - 7.9 \text{ kJ mol}^{-1} \quad (7.16)$$

Whereas Gibbs free energies of hydration for neutral solutes are often available in the literature, the values for charged species (conjugate acids or bases) are usually evaluated by means of a thermochemical cycle. The thermochemical cycles are shown in figures 7.1 and 7.2. Thus, taking pyridine (Py) and phenol (PhOH) as examples, $\Delta G°_{hyd}(PyH^+)$ and $\Delta G°_{hyd}(PhO^-)$ are given by eqs. 7.17 and 7.18, respectively,

$$\Delta G°_{ion}(PyH^+) = GB(Py) + \Delta G°_{hyd}(H^+) + \Delta G°_{hyd}(Py) - \Delta G°_{hyd}PyH^+) \quad (7.17)$$

$$\Delta G°_{ion}(PhO^-) = GB(PhOH) + \Delta G°_{hyd}(H^+) + \Delta G°_{hyd}(PhO^-) - \Delta G°_{hyd}PhOH) \quad (7.18)$$

Taking GB(Py) = 898.1 kJ mol^{-1}, [130]$\Delta G°_{hyd}(Py)$ = –11.7 kJ mol^{-1} for pyridine[112] and $pK_a(PyH^+)$ = 5.21[149] allows the Gibbs free energy of hydration of the pyridinium cation to be determined as $\Delta G°_{hyd}(PyH^+)$ = –256.1 kJ mol^{-1}. Similarly, taking GB(PhO^-) = 1432.0 kJ mol^{-1} ± 8.4 kJ mol^{-1}[127], $\Delta G°_{hyd}$(PhOH) = –19.7 kJ mol^{-1} for phenol[150] and pK_a(PhOH) = 10.02[151] allows the Gibbs free energy of hydration of the phenolate anion to be determined as $\Delta G°_{hyd}(PhO^-)$ = –255.7 kJ mol^{-1}. Both these values correspond to transfer of the ion from the gaseous state at 1 atmosphere to a 1 Mol dm^{-3} aqueous solution.

7.3) Methodology

The energies presented are the energies for the geometry-optimised structures of the compounds. The optimisation used symmetry where appropriate. All the structures were specified using Z-matrix coordinates. Dummy atoms included where three atoms were in a line, notably in the cyano-substituted systems. Calculations were also performed with explicitly represented waters, using dummy atoms within the dielectric continuum; however the results produced provided a substantially worse correlation than those without.

All the semiempirical calculated energies are based on a RHF semiempirical, AM1 and PM3, solute and all the DFT calculated energies were calculated using the BLYP/6-31+G(d) approach. The moderate BLYP/6-31+G(d) DFT approach was used as, at reasonable computational expense it effectively produced good results for the gaseous proton affinities and gas basicities, chapter 6. The aqueous enthalpies of formation for pyridines and phenols were calculated using MOPAC93 within the COSMO dielectric continuum. The aqueous Gibbs free energies of formation for pyridines and phenols were calculated using AMSOL6.1 with the SM5.4A and SM5.4P dielectric continuum models. The DFT aqueous Gibbs free energies of hydration and gas phase DFT Gibbs free energies of formation were used to calculate the Gibbs free energies of formation for pyridines and phenols using Gaussian 98 with CPCM representing solvation effects.

The keywords GNORM=0.0 (MOPAC), GCOMP=0.0 (AMSOL), KICK=3 (AMSOL), DDMIN=0.0, SCFCRT=1E-10 and HESS=3 were specified in all the semiempirical calculations. This increases the convergence criteria substantially over the

default. The eigenvector following optimiser[120-123] was used in MOPAC calculations wherever possible as this is highly recommended when using COSMO. [52]The eigenvector following optimiser led to structures of lower energy with substantially lower gradient norms than those produced using the default BGFS optimiser. Only results for the most stable conformers are reported here. A dielectric continuum of 78.4, the value for water, was used for all the COSMO calculations.

The density functional calculations were performed using following keywords were included in all the calculations, opt=(tight,GDIIS) freq=(NORAMAN) blyp/6-31+g(d) archive, and the following keywords were included when solvation effects were included CPCM[59], scrf=(cpcm,solvent=water). The opt=tight keyword was removed for systems that failed to converge notably for the nitro-substituted hydration calculations.

7.4) Results and Discussion

The ability of the theoretical semiempirical and density functional methods to reproduce experimental aqueous enthalpy and Gibbs free energy of deprotonation is to be evaluated. Regression analysis of plots of the calculated versus experimental values are summarised in Table 7.1. 30 phenolates and 35 pyridines were calculated using semiempirical methods and 7 phenolates and 7 pyridines were calculated using DFT methods. Due to the availability of experimental data, a small number of structures were compared to the experiment, as indicated in Table 7.1. Calculated aqueous COSMO enthalpies of deprotonation (eq. 7.17) are plotted versus experimental values for phenolates and pyridines in Figure 7.3 and 7.4 respectively. Semiempirically calculated aqueous Gibbs free energies of deprotonation (eq. 7.15) and DFT calculated aqueous Gibbs free energies of deprotonation (eq. 7.16) versus experimental values for phenolates and pyridines are shown in Figure 7.7 and 7.8 respectively. Where no experimental Gibbs free energies were available, their values were calculated from experimental pK_a values (eq. 7.18).

The enthalpies and Gibbs free energies of hydration of pyridines and pyridinium cations were investigated separately for the semiempirical methods. Semiempirically, AM1/COSMO and PM3/COSMO, calculated aqueous enthalpy of

hydration versus experimental values for pyridines and pyridinium ions are shown in Figure 7.5 and 7.6 respectively. Semiempirically, AM1/SM5.4A and PM3/SM5.4P calculated aqueous Gibbs free energies of hydration versus experimental values for pyridines and pyridinium ions are shown in Figure 7.9 and 7.10 respectively.

Table 7.1 – Regression analysis of experimental verses calculated enthalpies and Gibbs free energies of deprotonation, in kJ mol^{-1}.

Graph and Figure Number	Symbol	Class of compound	Method	Number of compounds	Gradient	Intercept	Correlation coefficient (r)	Root Mean Square error	Root Mean Square error in pK_a units
$\Delta H_{ion,calc}$ vs. $\Delta H_{ion,expt}$									
7.3	●	Phenols	AM1/COSMO	15	2.295	-44.4	0.517	20.1	3.5
7.3	■	Phenols	PM3/COSMO	15	1.943	-55.6	0.492	36.6	6.4
7.3	●	Phenols (No *x* -NO_2)	AM1/COSMO	12	1.141	-14.5	0.520	12.8	2.2
7.3	■	Phenols (No *x* -NO_2)	PM3/COSMO	12	0.956	-30.1	0.552	31.4	5.5
7.4	●	Pyridines	AM1/COSMO	22	0.648	-43.9	0.883	51.2	9.0
7.4	■	Pyridines	PM3/COSMO	22	0.330	-44.9	0.661	58.9	10.3
$\Delta G_{ion,calc}$ vs. $\Delta G_{ion,expt}$									
7.7	●	Phenols	AM1/SM5.4A	19	2.225	-45.1	0.820	22.0	3.8
7.7	■	Phenols	PM3/SM5.4P	19	2.221	-64.5	0.801	11.8	2.1
7.7	□	Phenols	BLYP/631+G(d)/CPCM	6	-0.158	28.7	0.084	32.0	5.6
7.7	●	Phenols (No *x* -NO_2)	AM1/SM5.4A	16	1.190	12.0	0.683	23.0	4.0
7.7	■	Phenols (No *x* -NO_2)	PM3/SM5.4P	16	1.055	-0.1	0.607	6.7	1.2
7.7	□	Phenols (No *x* -NO_2)	BLYP/631+G(d)/CPCM	5	0.833	-24.5	0.332	34.8	6.1
7.8	●	Pyridines	AM1/SM5.4A	28	0.791	-48.7	0.873	43.2	7.6
7.8	■	Pyridines	PM3/SM5.4P	28	0.612	-55.3	0.766	58.7	10.3
7.8	□	Pyridines	BLYP/631+G(d)/CPCM	5	0.274	8.6	0.542	22.5	3.9

Table 7.2 – Regression analysis aqueous hydration enthalpies and Gibbs free energies for pyridines and pyridinium ions, in kJ mol^{-1}.

Graph and Figure Number	Symbol	Class of compound	Method	Number of compounds	Gradient	Intercept	Correlation coefficient	Root Mean Square error
$\Delta H_{hyd,calc}$ vs. $\Delta H_{hyd,expt}$								
7.5	◆	Pyridines	AM1/COSMO	18	0.383	-19.3	0.592	18.5
7.5	■	Pyridines	PM3/COSMO	18	0.327	-16.9	0.472	23.8
7.6	◆	Pyridinium ions	AM1/COSMO	20	1.243	22.5	0.899	34.6
7.6 (with outlying points)	◆	Pyridinium ions	AM1/COSMO	22	1.028	-25.7	0.637	39.2
7.6	■	Pyridinium ions	PM3/COSMO	20	1.439	59.4	0.907	42.6
7.6 (with outlying points)	■	Pyridinium ions	PM3/COSMO	22	1.324	32.3	0.636	51.6
$\Delta G_{hyd,calc}$ vs. $\Delta G_{hyd,expt}$								
7.9	◆	Pyridines	AM1/SM5.4A	6	0.825	-3.3	0.729	11.5
7.9	■	Pyridines	PM3/SM5.4P	6	0.221	-3.9	0.262	7.8
7.10	◆	Pyridinium ions	AM1/SM5.4A	6	1.220	15.7	0.775	14.0
7.10	■	Pyridinium ions	PM3/SM5.4P	6	1.103	7.2	0.804	6.0

7.4.1) Enthalpy of deprotonation

The semiempirical COSMO calculated enthalpies of deprotonation have been plotted against experimental enthalpies of deprotonation for 15 substituted phenols[150, 152-155] and substituted pyridines, [156]shown in Figures 7.3 and 7.4 respectively. The enthalpies of deprotonation follow similar trends to the PA, chapter 6, although the calculated enthalpies of deprotonation are considerably less accurate than PA. This indicates there is considerably more error in the treatment of solvation than in the underlying description of the solute particularly for the DFT approaches.

AM1/COSMO reproduces the experimental enthalpies of deprotonation significantly better than PM3/COSMO with the correlation coefficient and gradient closer to unity, shown in Table 7.1. The sensitivity to the effect of the substituent, denoted by the gradient, is significantly overestimated for phenolates with gradients of 2.295 (AM1/COSMO) and 1.943 (PM3/COSMO) and underestimated for pyridines with gradients of 0.648 (AM1/COSMO) and 0.330 (PM3/COSMO). When the outlying nitro-phenolate substituents were removed, the gradient was much improved, giving gradients of 1.141 and 0.956 for AM1/COSMO and PM3/COSMO respectively. This and the subsequent anomalous results for nitro-substituted phenolates may be due to several factors, for example the neglect of explicit solvation, which is particularly important for the nitro-substituted compounds. There is also weakness of PM3 parameterization for compounds containing nitrogen atom. [157] The correlation coefficients are lower than in the PAs and closer to unity for the pyridines, with values of 0.883 (AM1/COSMO) and 0.661 (PM3/COSMO) compared to the phenolates with values of 0.517 (AM1/COSMO) and 0.492 (PM3/COSMO). The exclusion of the nitro-phenolates substituents improved the correlation coefficient, but not significantly.

The absolute values of the calculated enthalpies of deprotonation are systematically lower than experimental values, particularly for the pyridines. This leads to high RMS errors with all the AM1/COSMO calculations having lower RMS errors than the PM3 calculations. The pyridine systems especially have high RMS values with values of 51.2 and 58.9 kJ mol^{-1} for AM1/COSMO and PM3/COSMO respectively, showing a similar trend to the PA results. The RMS error for the phenolates of 20.1 (AM1/COSMO) and 36.6 (PM3/COSMO) kJ mol^{-1} was

improved to 12.8 (AM1/COSMO) and 31.4 (PM3/COSMO) kJ mol^{-1} when the nitro-phenol substituents were excluded.

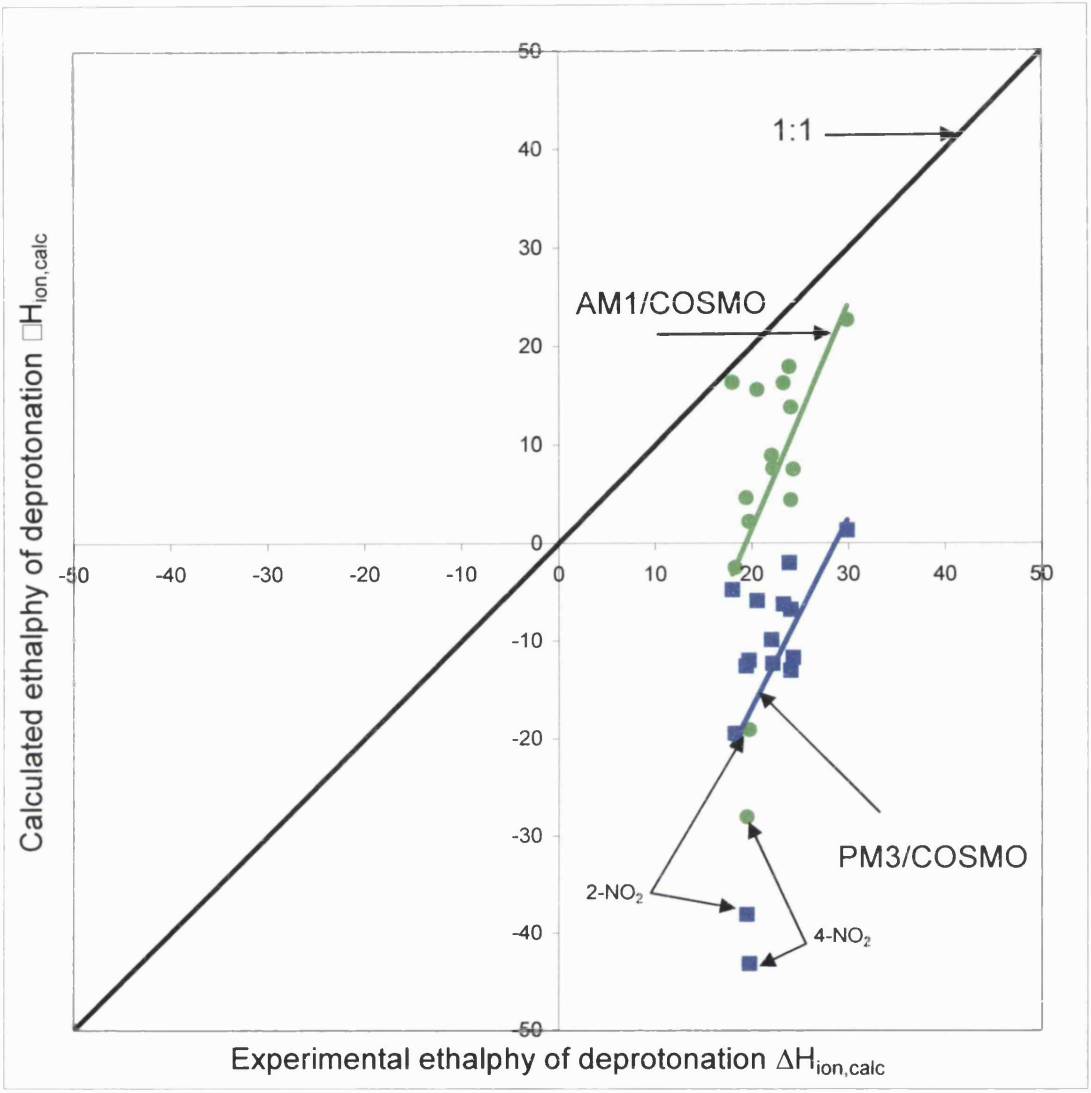


Figure 7.3 – Calculated AM1/COSMO and PM3/COSMO for enthalpies of deprotonation for 15 substituted phenolate anions versus experimental enthalpies of deprotonation, in kJ mol^{-1}. The most significan outliers on the chart are the nitro-substituted compounds.

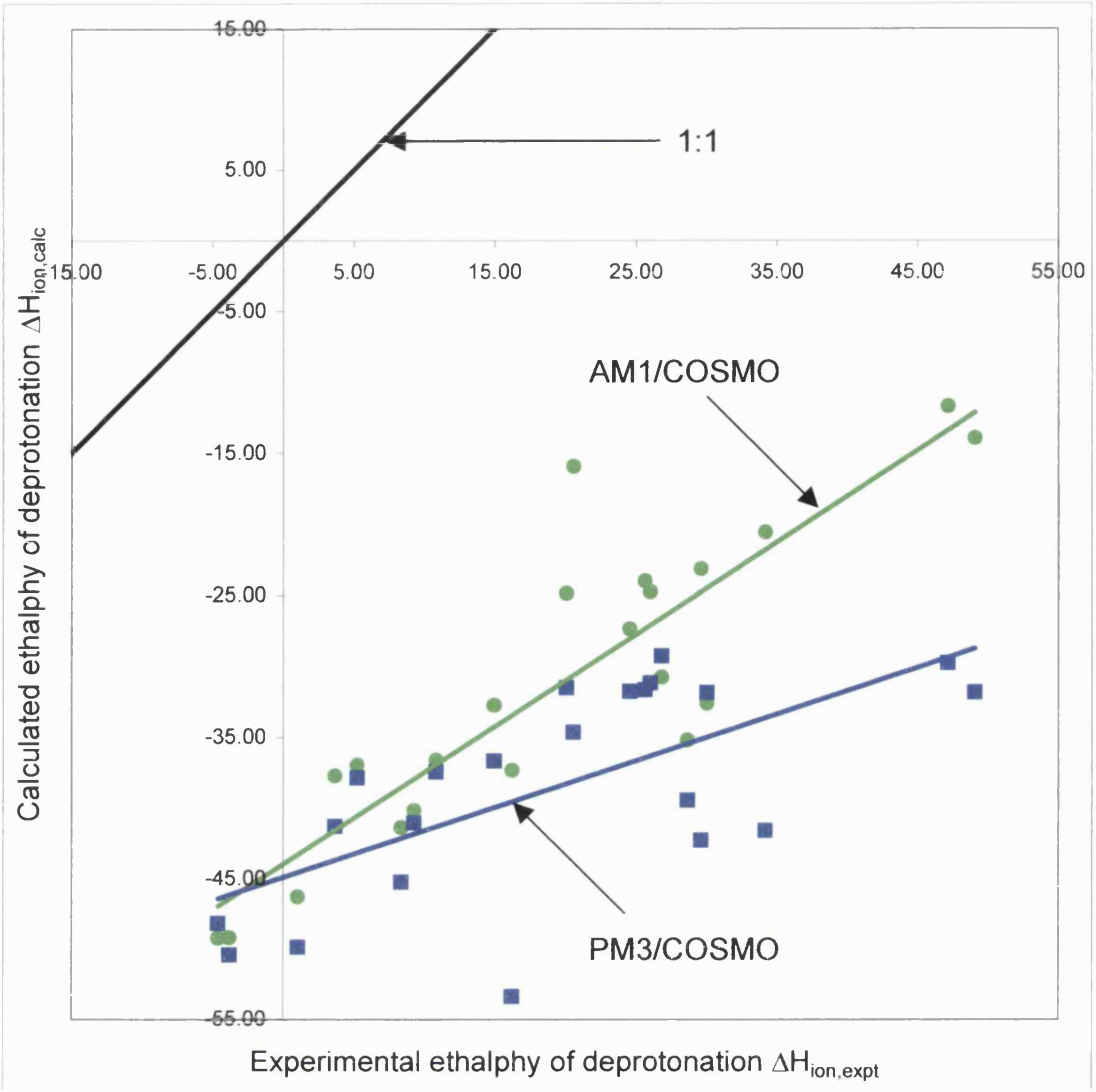


Figure 7.4 – Calculated AM1/COSMO and PM3/COSMO for enthalpies of deprotonation for 22 substituted pyridines versus experimental enthalpies of deprotonation, in kJ mol^{-1}.

7.4.2) Enthalpies of hydration of pyridines and pyridinium cations

Two components of the enthalpy of deprotonation, the enthalpies of hydration for the pyridine systems, which had the highest RMS errors for the enthalpies of deprotonation were examined further. The calculated enthalpy of hydration of 18 pyridines and 20 pyridinium cations were plotted against experimental values, shown in Figure 7.5 and 7.6 with regression analysis summarised in Table 7.2. Fewer points are plotted than for the enthalpies of deprotonation, due to the lack of experimental data. The selection of compounds has a significant effect on the graphs and therefore the regression analysis is given. There are two notable outliers, $4NO_2pyH^+$ and $4N(CH_3)_2pyH^+$, on the plot of calculated versus experimental enthalpies of hydration. The correlation to the line of best fit0 improves markedly without these two outliers present with the correlation coefficient changing from 0.637 (AM1/COSMO) and 0.637 (PM3/COSMO) to 0.899 (AM1/COSMO) and 0.907 (PM3/COSMO). The neutral pyridines had fewer and less significant outliers.

The neutral pyridines showed a considerable under sensitivity to the effect of the substituent with a very low gradient of 0.383 (AM1/COSMO) and 0.327 (PM3/COSMO). The calculations of the enthalpy of hydration for the pyridinium cations are also poor; these were over sensitive to the effect of the substituent, although they are closer to one than the neutral pyridines with gradients of 1.243 (AM1/COSMO) and 1.439 (PM3/COSMO) when the outliers were excluded; the inclusion of the two outliers that reduce the reliability of the line of best fit reduces the gradient towards unity. The RMS error is comparatively large, around the same magnitude as in the enthalpies of deprotonation, ranging from 18.5 to 23.8 kJ mol^{-1} and 34.6 to 51.6 kJ mol^{-1} for the pyridines and pyridinium cations respectively. This high RMS error is due to significant deviation of the calculations from the absolute values of the enthalpies of hydration, as the calculations systematically overestimate the hydration of neutral pyridines and systematically underestimate the hydration pyridinium cations.

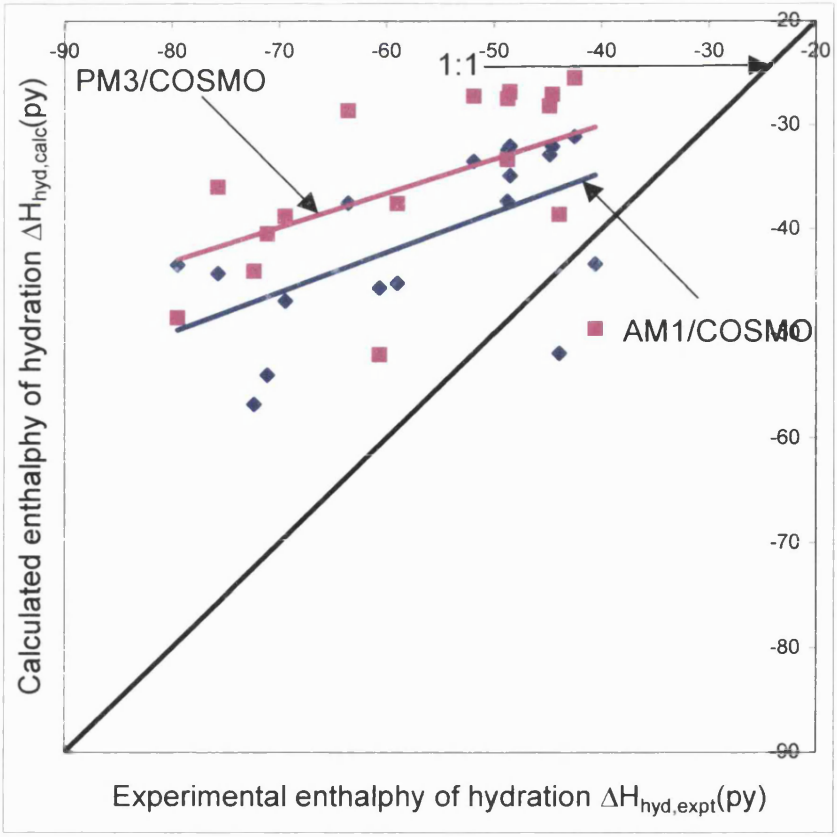


Figure 7.5 – Calculated semiempirical enthalpy for hydration of 18 substituted pyridines versus experimental[114, 158] in kJ mol^{-1}.

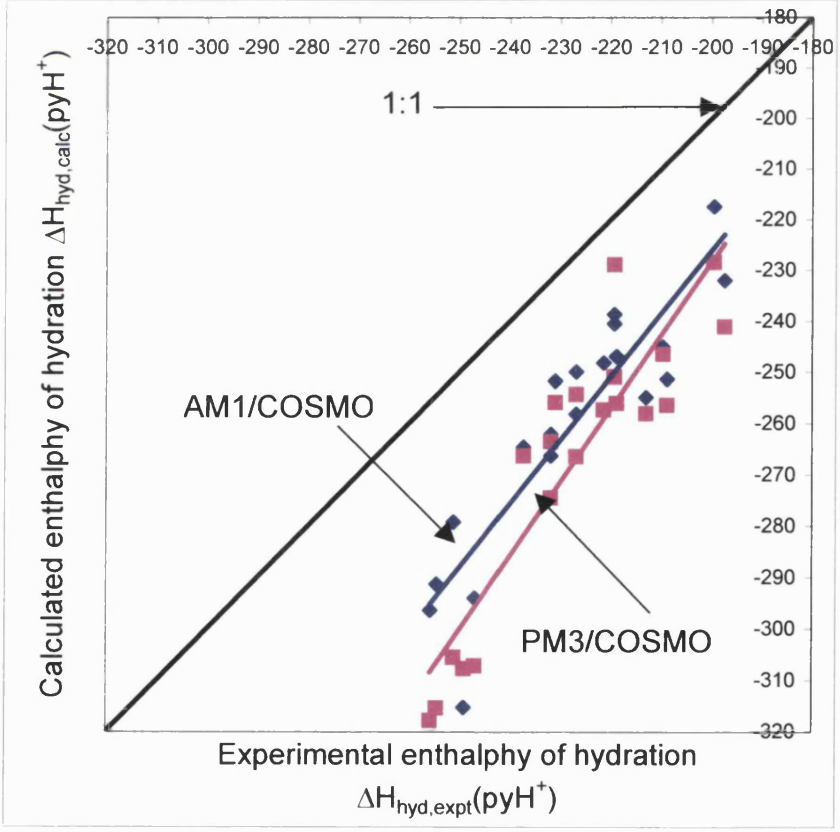


Figure 7.6 – Calculated semiempirical enthalpy for hydration of 20 substituted pyridinium ions versus experimental[114] in kJ mol^{-1}.

7.4.3) Gibbs free energy of deprotonation

Semiempirical SM5.4 and DFT/CPCM Gibbs free energies of deprotonation have been plotted against experimental Gibbs free energies of deprotonation for 19 substituted phenols[150-155, 159] and 28 substituted pyridines[149, 159] and are shown in Figures 7.7 and 7.8 respectively. These generally follow similar trends in the regression analysis to that of the PA and enthalpy of deprotonation. Semiempirical SM5.4*x* calculations reproduced the experimental Gibbs free energies of deprotonation better than semiempirical COSMO reproduced experimental enthalpies of deprotonation. The reason that SM5.4*x* is better is probably solely due to the different experimental points used for comparison. The regression analysis has been shown to be highly dependent upon which points are plotted. DFT/CPCM did not reproduce experimental Gibbs free energies as well as semiempirical SM5.4*x* approaches.

The calculated absolute values of the Gibbs free energy are systematically lower than experimental values for the semiempirical pyridines. The DFT calculated absolute Gibbs free energies were also systematically lower for both phenolates and pyridines. This systematic difference in the absolute values leads to large RMS errors, particularly for the pyridines. The sensitivity to the effect of the substituent, shown in the gradient, is consistently overestimated by semiempirical methods for phenolates with values of 2.225 (AM1/SM5.4A) and 2.221 (PM3/SM5.4P) and underestimated for the pyridines with gradients of 0.791 (AM1/SM5.4A) and 0.791 (PM3/SM5.4P). The semiempirical phenolates gave a much closer gradient to one when the outlying nitro-substituents were excluded, giving values of 1.190 (AM1/SM5.4A) and 1.055 (PM3/SM5.4P). The DFT/CPCM approach gave an extremely poor gradient of –0.158 for the phenolates, which improved to 0.833 when 4-nitrophenolate was excluded, and a very poor gradient of 0.276 for the pyridines. Anomalous energies for the nitro-phenolates are seen throughout the solvation systems, yet it is not seen in the gas phase calculations; therefore this is likely to be due to an incorrect treatment of solvation; this probably due to the neglect of specific solvation to the nitro-substituents. The correlation coefficient for the semiempirical best-fit lines ranged from 0.607 to 0.820 for the phenolates and 0.766 to 0.873 for the pyridines. The correlation coefficient for DFT/CPCM was 0.332 for the phenolates (when 4-nitrophenolate was excluded) and 0.542 for the pyridines, these add doubt to the reliability of the graphs.

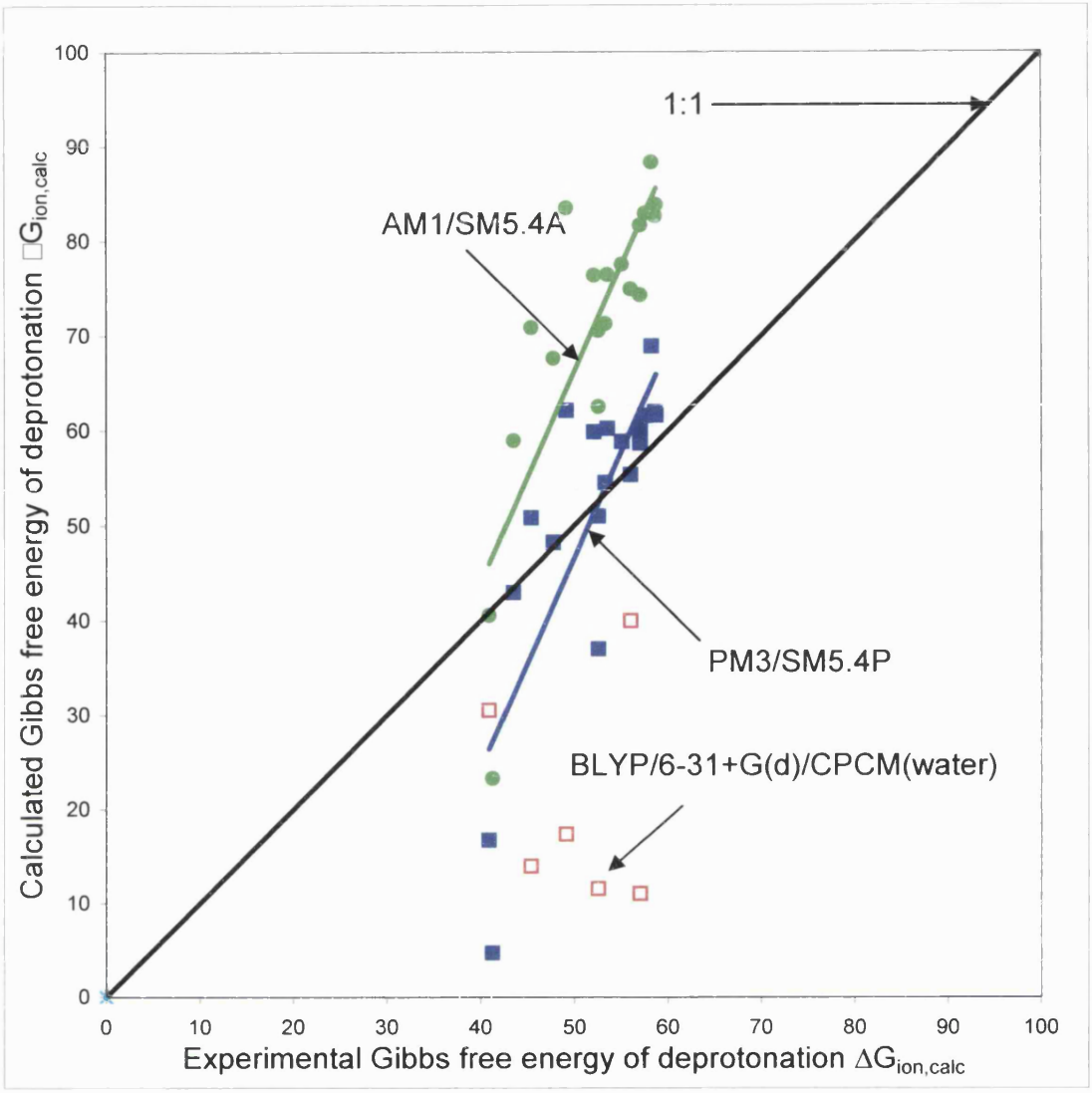


Figure 7.7 – Calculated AM1/SM5.4A and PM3/SM5.4P for Gibbs free energy of deprotonation for 19 substituted phenolate anions versus experimental Gibbs free energy of deprotonation and BLYP/6-31+(d)/CPCM(water) calculated Gibbs free energy of deprotonation for 6 substituted phenolate anions versus experimental Gibbs free energy of deprotonation, in kJ mol^{-1}. The line of best fit is omitted from the results of the BLYP6-31+(d)/CPCM(water) calculations as there is not a good enough correlation between the points to produce a meaningful line.

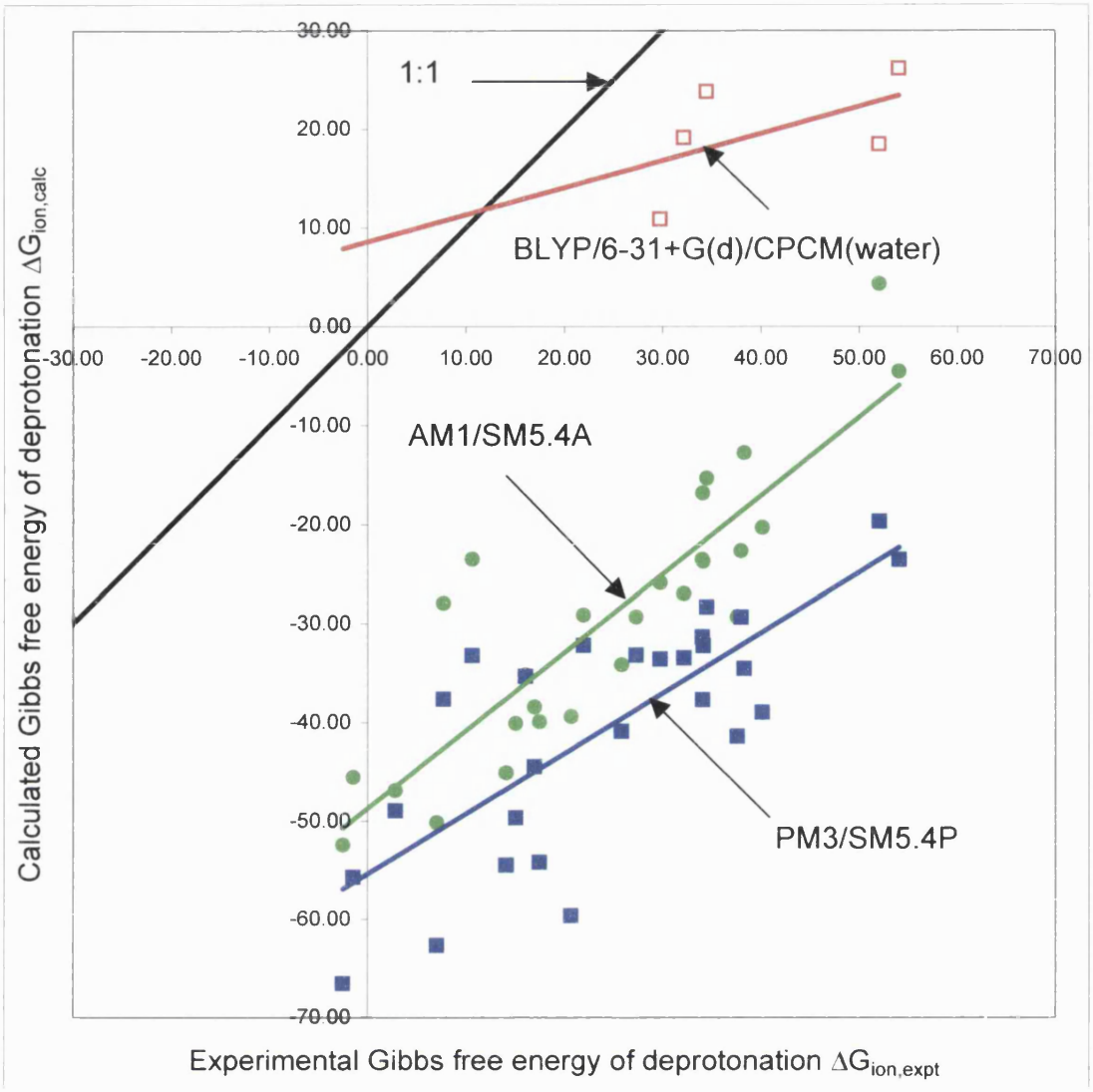


Figure 7.8 – Calculated AM1/SM5.4A and PM3/SM5.4P for Gibbs free energy of deprotonation for 28 substituted pyridines versus experimental Gibbs free energy of deprotonation and BLYP/6-31+(d)/CPCM(water) calculated Gibbs free energy of deprotonation for 6 substituted pyridines versus experimental Gibbs free energy of deprotonation, in kJ mol^{-1}.

7.4.4) Gibbs free energy of hydration of pyridines and pyridinium cations

The calculated Gibbs free energies of hydration for the 6 pyridine and 6 pyridinium cationic species were plotted against experimental values, Figures 6.9 and 6.10 with a summary of regression analysis in Table 8.2. The significance of the results is reduced, due to the limited amount of experimental Gibbs free energies of hydration with which to compare. The pyridinium cations have near unity, 0.998 (AM1/SM5.4A) and 0.982 (PM3/SM5.4P), correlations coefficients showing an excellent correlation to the line of best fit. The correlation coefficients are comparatively poor for the neutral pyridines, especially for the AM1/SM5.4A method giving correlation coefficients of 0.727 (AM1/SM5.4A) and 0.274 (PM3/SM5.4P). This casts considerable doubt on the best-fit line for the PM3/SM5.4P neutral pyridines calculations.

The sensitivity to the effect of the substituent, shown by the gradient, was overestimated for the pyridinium cations with gradients of 1.556 (AM1/SM5.4A) and 1.348 (PM3/SM5.4P) and significantly underestimated for the neutral pyridines with gradients of 0.727 (AM1/SM5.4A) and 0.274 (PM3/SM5.4P). The low correlation coefficient for the neutral pyridines using PM3/SM5.4P raises doubt as to the accuracy of this very poor gradient, and these combined raise doubts over the accuracy of this method for calculating the Gibbs free energy of hydration of neutral pyridines. The absolute values of the Gibbs free energies of hydration for the pyridines and pyridinium ions are close to experimental values. This is shown in comparatively low RMS errors of 7.8 (AM1/SM5.4A) and 11.5 (PM3/SM5.4P) kJ mol^{-1} for the neutral pyridines and 14.0 (AM1/SM5.4A) and 6.0 (PM3/SM5.4P) kJ mol^{-1}. The intercepts are highly gradient dependent, as the points are clustered in a comparatively small Section of the graph, therefore as with the other regression analysis the significance of the intercept is suspect. Overall the pyridinium cation is represented better than neutral pyridines by all the computational approaches and the error is more prevalent in the neutral pyridines. This conclusion is unexpected, as there are fewer ions in the parameterisation sets of the solvation models than neutrals.

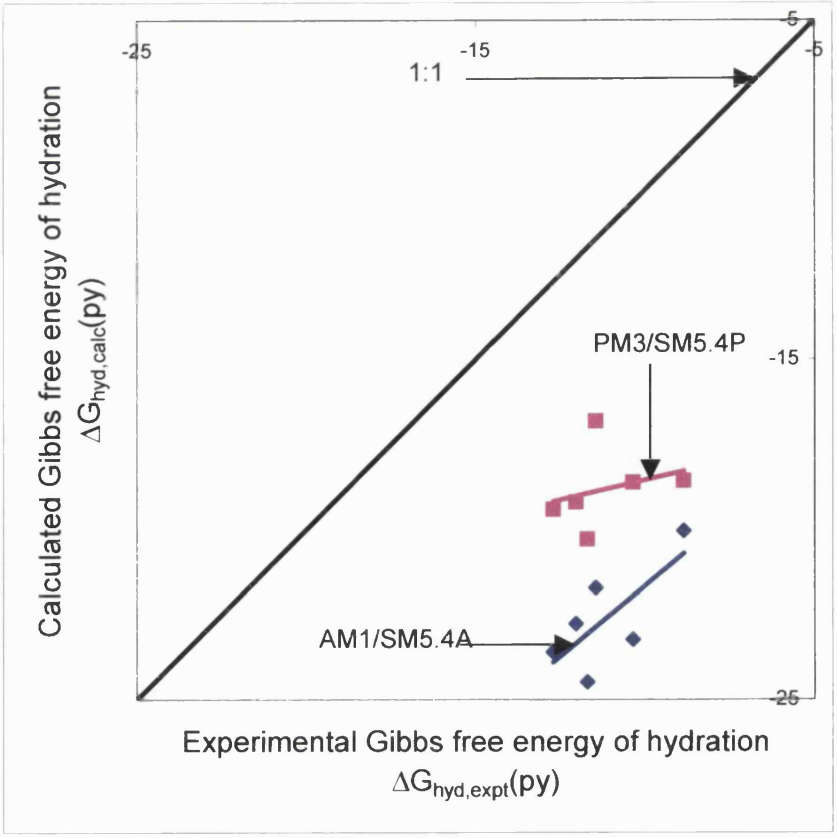


Figure 7.9 – Calculated semiempirical Gibbs free energy of hydration of 6 substituted pyridines versus experimental[114, 158] in kJ mol^{-1}.

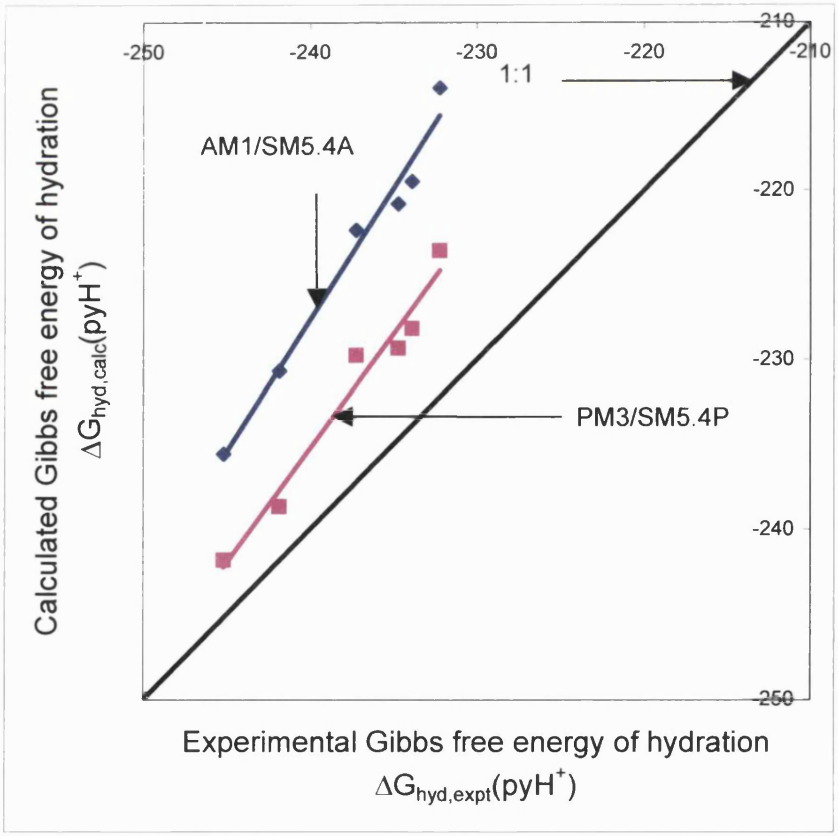


Figure 7.10 – Calculated semiempirical Gibbs free energy of hydration of 6 substituted pyridinium ions versus experimental[114] in kJ mol^{-1}.

7.5) Conclusion

The computational reproduction of aqueous phase energetics of deprotonation of pyridines and phenolates was substantially poorer than in the gaseous phase, chapter 6. The majority of the errors are therefore in the solvation treatment of the aqueous enthalpy and Gibbs free energy of deprotonation not in the underlying solute representation. However, the aqueous energetics (hydration and deprotonation) approximately followed the same trends overall as the gaseous energetics (formation and deprotonation); the aqueous Density Functional Theory (DFT) were extremely poor in comparison. The solvated deprotonation calculations are systematically under sensitive to the effect of the substituent for the pyridines and systematically overestimate this effect for the phenolates. It is indicated in the hydration energies that this under sensitivity is derived from the solvation treatment of the neutral pyridines; where the substituent effect is underestimated for the neutral pyridines to a greater degree than it is overestimated for the pyridinium cation. The errors in the substituent sensitivity are substantially greater in the deprotonation energies, gradients of plots of calculated versus experimental being further from one, than in the gaseous phase.

The RMS error of the computational methods compared to the experimental values for both deprotonation and hydration energetics vary greatly depending on the combination of molecules studied and method used. The systematic error in the sensitivity to the effect of the substituent increases this difference between calculated and experimental absolute energies giving a larger RMS error. By a judicial selection of 'well behaved' points (substituents) it is possible to get a low RMS, as is prevalent in published studies. The assessment of the accuracy of computational methods necessitates both a large set of substituents and the consideration of both correlation and substituent sensitivity as well as the RMS. The published studies also quote the RMS error in pK_a units, a large unit which also appears to reduce the errors quoted (RMS is tabulated both in kJ mol^{-1} and pK_a units in Table 7.1). The exclusion or inclusion of a single substituent has been shown to affect the correlation to experimental values markedly. The nitro-substituents were outliers on the graphs plotted; without these a much better correlation between calculated and experimental values was found. It is worth noting that nitro-phenol was the substituent chosen, for good experimental

reasons, for the study of triazinyl reactivity by A. Williams and co-workers[92] the reproducibility of these experiments computationally is studied in chapter 5. As the nitro-substituent was an outlier for the aqueous phase (not gas phase – chapter 6) calculations, this casts doubt on the ability of continuum solvation models to model triazine reactivity. The reason why the nitro-substituent was described particularly badly using continuum approaches is due to the need to model explicit solvation to the nitro groups.

Overall AM1 (AM1/COSMO and AM1/SM5.4A) produced the most consistent energies of deprotonation and hydration with the highest correlation coefficient, followed closely by PM3 (PM3/COSMO and PM3/SM5.4P) with DFT (BLYP/6-31+G(d)/CPCM(water)), showing little consistency in values or to the experimental values. This is a large change from the gas phase where DFT performed the best and showed the most consistent agreement to experiment by far. Therefore the representation of the solvent has the largest error and more accurate intensive treatments may be required. The semiempirical methods have been shown to perform better in the presence of a continuum than density functional approaches. This may be as these have been implemented sooner and are more refined and calibrated than the density functional approaches. The extra parameterisation used in these semiempirical calculations may be more suitable to be adapted and fine tuned.

7.6) Evaluation

The computational representation of solvation energetics using continuum approaches for small nucleophiles has considerably more errors than in the description of the solute. The semiempirical solvation models reproduced experimental energetics better than the more computationally expensive DFT equivalent. The best correlation to experimental energetics of the methods assessed was shown using the semiempirical AM1 parameterisation combined with the SM5.4A solvation model.

Chapter 8

Crystal structure geometries of triazines and trinitrobenzenes

Are density functional methods able to reproduce crystal structure geometries?

8.1) Introduction

The ability of theoretical methods to reproduce molecular structures is of considerable importance to computational chemists. This can be assessed by comparing computationally derived structures to experimental crystal structures. Extensive experimental studies of heterocyclic chemistry have been made and a wide variety of texts and publications in the area of heterocyclic chemistry, are available, notably by Katritzky and co-workers. [160-162]These have shown that 1,3,5-triazine has several physical properties expected for aromaticity, such as diamagnetic anisotropy and a ring-current contribution to average diamagnetism. However, bond lengths do not favour this physical aromaticity. These structural physical properties and corresponding behaviour make the structures of 1,3,5-triazinyl adduct species of particular interest.

The experimental structures need to be of a sufficient standard, having a low R-factor, preferably below 0.1, to allow direct comparison with computationally derived structures. Triazine and trinitrobenzene containing structures, similar to the substrate examined in Chapter 5 from the Cambridge Structural Database[163] (CSD), have been compared to calculated density functional theory (BLYP and B3LYP[109, 110]) derived structures. Due to crystal packing effects in the CSD structures this is not ideal, however these provide good reference structures and should provide a meaningful test of the computational methods, especially for the core of heavy atoms.

8.2) Cambridge Structural Database search

The CSD was searched for structures containing triazine and trinitrobenzene rings, firstly with (a) monosubsituted 'reactant' structure and (b) with two T atoms, where T can be C, N, O, F, P, S, Cl, Se, Br, I or At, attached to the same carbon in the ring, Figure 8.1. The bond lengths in the ring were assigned B1 to B6 clockwise away from the tetrahedral carbon and the two bond lengths to T were denoted R1 and R2. The angle between the two T atoms was denoted as A1 and the angles and torsional angles with the ring were denoted A2/A3 and T1/T2 respectively. Amidines are a further group which could be compared to triazine; there was, however, no comparable data for these structures in the CSD.

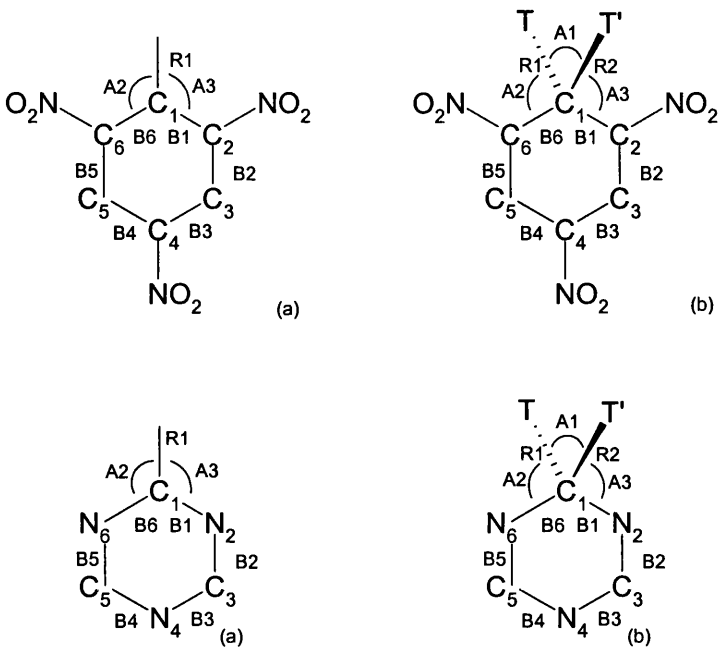


Figure 8.1 – Schematic of the core of the (a) monosubstituted and (b) symmetrical disubstituted trinitrobenzene and triazine species searched for on the CSD.

The group of compounds were then narrowed so only compounds with Meisenheimer type properties were present, shown in Figure 8.3. The full CSD information, reference, formula, compound name, crystal information and structure are given in Appendix A. Meisenheimer complexes are geminally disubstituted with electronegative heteroatom groups. That is to say negatively charged systems with a positive counter ion, which contain no covalent bonds to triazinyl nitrogens or to nitro groups on the trinitrobenzene ring. Meisenheimer complexes[7] are generally formed by attack at the position para to at least one

electronegative group (2,4,6- for triazine and 1,3,5- for TNB), as shown for the classical Meisenheimer reaction in Figure 8.2, as this causes the greatest exothermic change. [89]

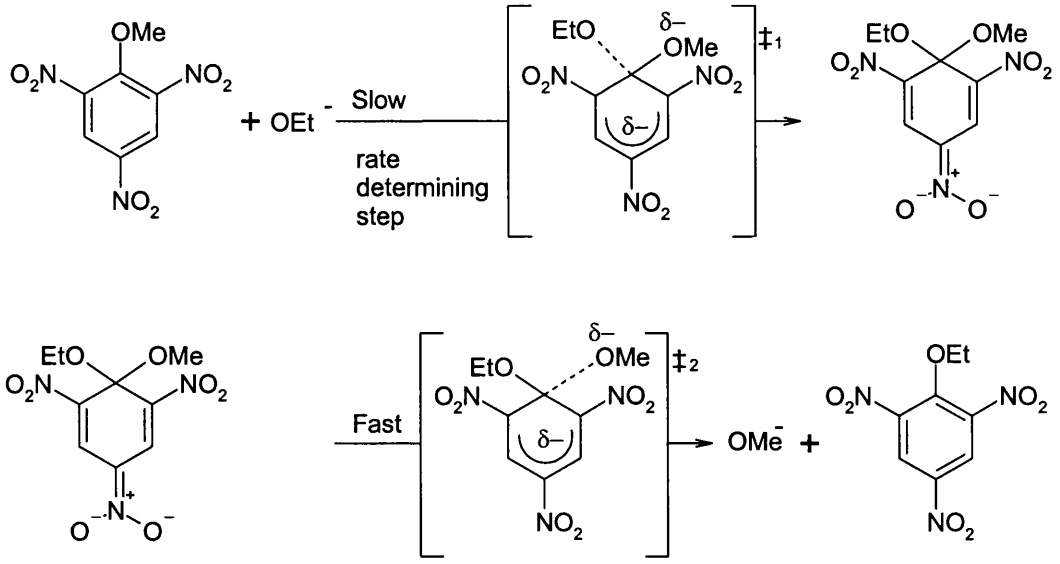
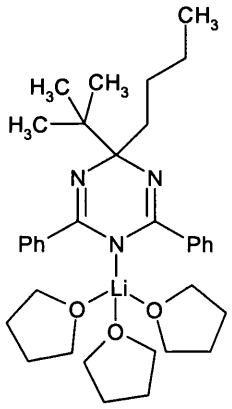
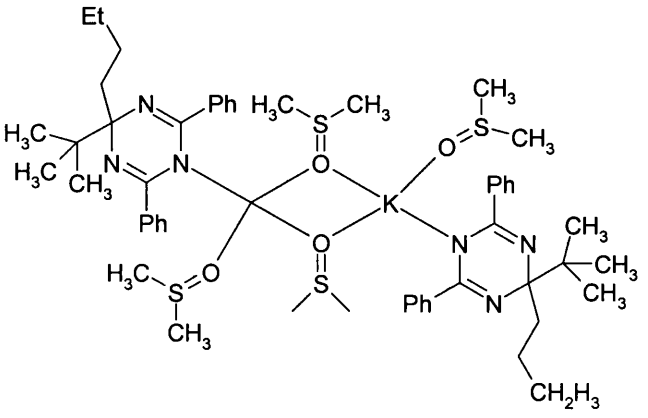


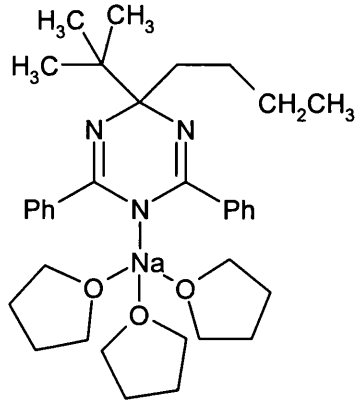
Figure 8.2 – The classical Meisenheimer reaction mechanism[7] of ethyl picrate with a methoxide anion. (as shown in the introductory Section)



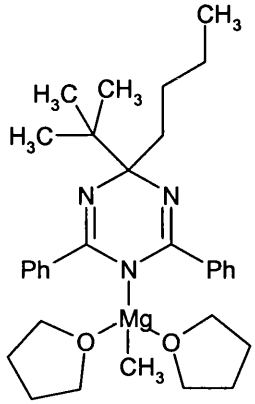
LAFBUP10



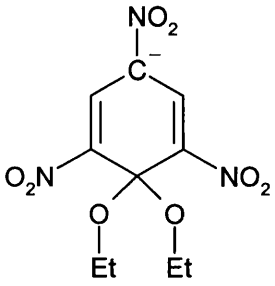
REMGOF



REMGUL

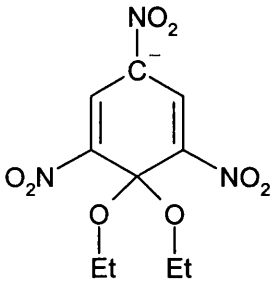


YIZGIX



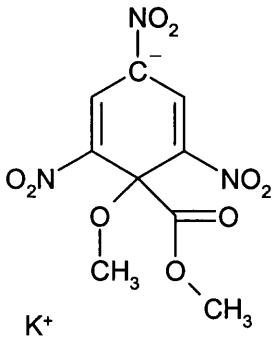
K^+

KDMTNB10



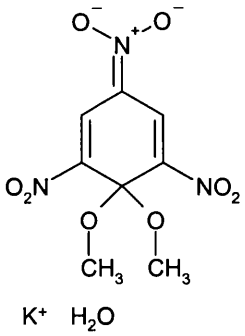
Cs^+

MNXCHY



K^+

TNPCSE



K^+ H_2O

TNPTKE

Figure 8.3 – CSD reference and the corresponding structure. The reference in italics is the most representative structure chosen.

8.3) Analysis of Cambridge Structural Database search

Bond lengths, B2 and B3 in Table 8.4 and 8.5, in the triazine and trinitrobenzene rings were plotted against one another; relatively short and equal bond lengths would indicate a delocalised ring whereas, significantly different bond lengths would denote a relatively localised system. All the triazine B2 versus B3 plots, Figure 8.4, show that the majority of the structures from the CSD have different bond lengths, over 0.075Å different, indicating a relatively localised ring system. These relatively localised systems had a longer B3 bond than B2 suggesting that B2 had more π-bond character. The other bond lengths within the ring upon further investigation, by plotting of the remaining ring bond lengths against each other, are not equal throughout. The trinitrobenzene structures B2 versus B3 plots, Figure 8.5, show only one compound (TNPCSE - 2,4,6-Trinitrophenolate-cesium ethoxide complex) having equal bond lengths, which indicates a relatively delocalised system. Both structures within the unit cell exhibited this property. The remaining structures, KDMTNB10 and TNPTKE, had B3 longer than B2, suggesting B2 is the double bond and KMXCHY had a longer B2 than B3, suggesting B3 is a double bond in this structure.

Table 8.3 - Bond lengths (in Å) for 5 CSD entities and 6 observable disubstituted triazines. The structure in italics is the most representative structure, the structure closest to the calculated structure.

CSD Reference	B1	B2	B3	B4	B5	B6	R1	R2	RFAC
LAFBUP10	*1.457*	*1.287*	*1.386*	*1.371*	*1.291*	*1.476*	*1.533*	*1.561*	*0.092*
REMGOF	1.484	1.298	1.372	1.367	1.293	1.475	1.537	1.583	0.046
REMGUL 1	1.478	1.288	1.375	1.366	1.286	1.475	1.538	1.577	0.095
REMGUL 2	1.474	1.292	1.376	1.366	1.297	1.483	1.534	1.567	0.095
YIZGIX	1.480	1.280	1.389	1.397	1.279	1.473	1.474	1.579	0.057
Mean	1.475	1.289	1.380	1.373	1.289	1.476	1.523	1.573	
Standard deviation	0.010	0.007	0.007	0.013	0.007	0.004	0.028	0.009	
Minimum	1.457	1.280	1.372	1.366	1.279	1.473	1.474	1.561	
Maximum	1.484	1.298	1.389	1.397	1.297	1.483	1.538	1.583	

Table 8.4 - Bond lengths (in Å) for 4 CSD entities and 6 observable disubstituted TNBs. The structure in italics is the most representative structure, the structure closest to the calculated structure.

CSD Reference	B1	B2	B3	B4	B5	B6	R1	R2	RFAC
KDMTNB10	*1.475*	*1.353*	*1.397*	*1.391*	*1.358*	*1.516*	*1.430*	*1.422*	*0.130*
KMXCHY	1.511	1.429	1.381	1.417	1.389	1.427	1.574	1.424	0.150
TNPCSE 1	1.495	1.359	1.371	1.448	1.312	1.507	1.395	1.434	0.102
TNPCSE 2	1.509	1.384	1.394	1.424	1.301	1.563	1.407	1.390	0.102
TNPTKE 1	1.516	1.354	1.404	1.406	1.347	1.512	1.425	1.410	0.064
TNPTKE 2	1.512	1.335	1.414	1.406	1.352	1.515	1.416	1.417	0.064
Mean	1.503	1.369	1.394	1.415	1.343	1.507	1.441	1.416	
Standard deviation	0.015	0.033	0.016	0.020	0.032	0.044	0.066	0.015	
Minimum	1.475	1.335	1.371	1.391	1.301	1.427	1.395	1.390	
Maximum	1.516	1.429	1.414	1.448	1.389	1.563	1.574	1.434	

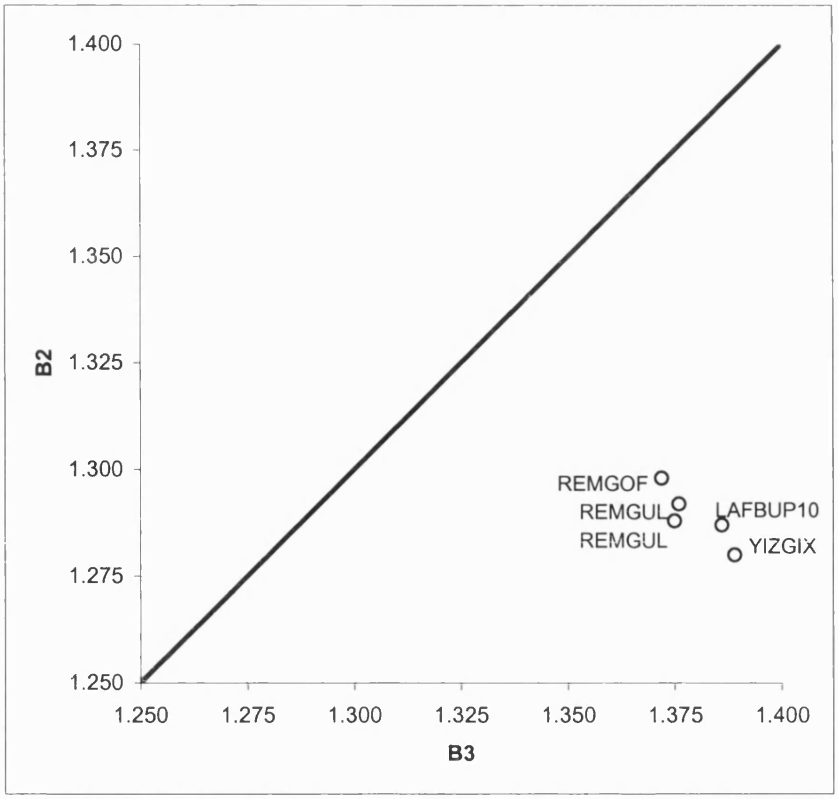


Figure 8.4 – Plot of B2 versus B3 for the CSD triazine structures in Å

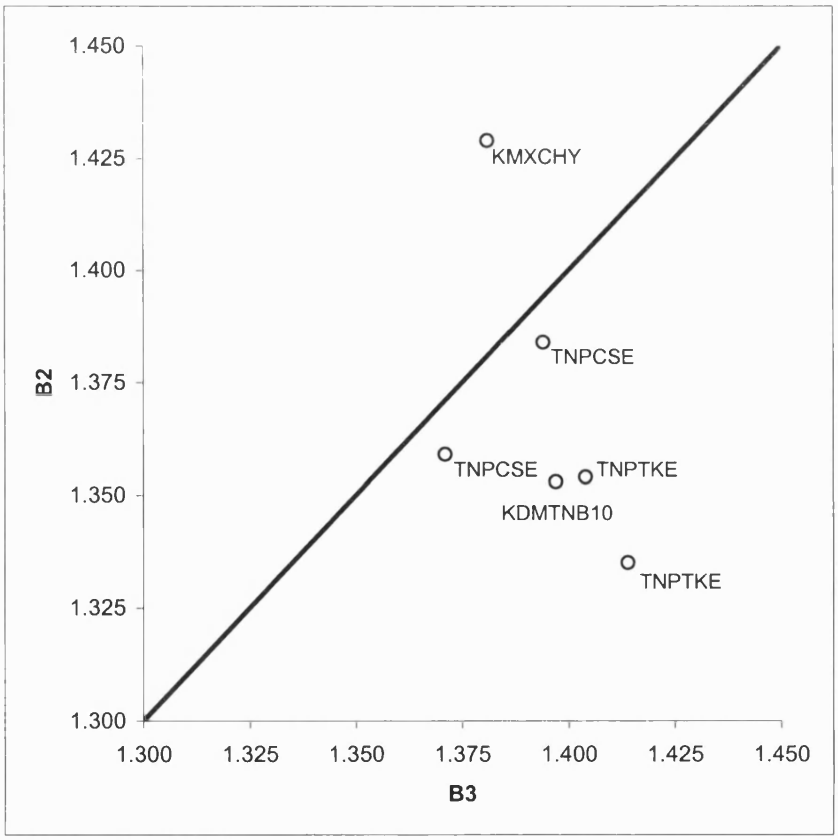


Figure 8.5 – Plot of B2 versus B3 for the CSD TNB structures in Å

Bond lengths, R1 and R2, to the nucleophile/leaving group, denoted T, in the triazine and TNB structures were plotted in order to assess how the nucleophile/leaving group bind to the crystallographic structures. All the bond lengths were longer than for those in the monosubstituted structures as expected, due to increased steric hindrance, for both triazine and TNB structures. The triazinyl structures had similar R1 and R2 bond lengths, Figure 8.6, indicating structures similar to classical Meisenheimer intermediates, with four entities within 0.02Å and only one 0.05Å from one another. The YIZGIX structure, the outlier, has n-butyl and t-butyl substituents the bond length to the sterically hindered t-butyl substituent is longer than the n-butyl substituent. The extent of this difference is also strongly influenced by the Mg^{2+} ion that pulls electrons from the ring and in turn shortens the bond to the electron donating aryl-substituent, the effect is less pronounced on the sterically hindered t-butyl substituent. The disubstituted TNB structures, Figure 8.7, have similar bond lengths with five out of six structures within 0.04Å, the exception being KMXCHY which differs by over 0.12Å. This is caused by the difference in the substituents bonding atom, with KMXCHY having a C-O and C-C bond, whereas the other entities have the same type of bond to nucleophiles and leaving groups.

The majority of the structures found have different B2 and B3 bond lengths, indicating relatively localised ring systems. The favoured resonance canonical is related to the position of the cation, with B2 and B5 generally the double bonds for the structures investigated. The bond lengths to the substituents on the tetrahedral carbon, R1 and R2, were approximately the same: that indicates a similarity to classical Meisenheimer intermediates.

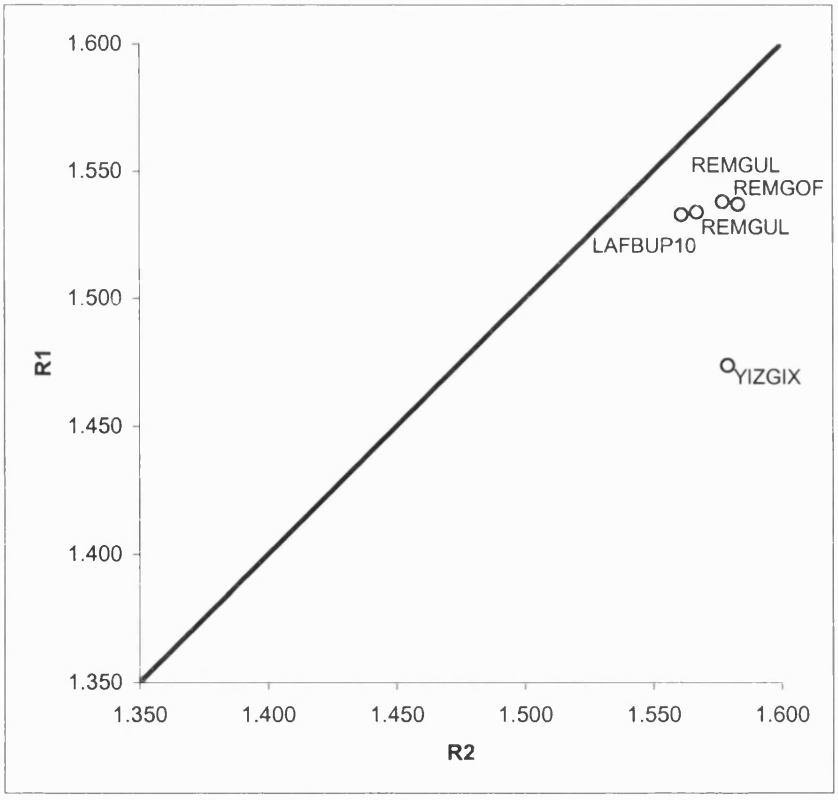


Figure 8.6 – Plot of R1 versus R2 for the CSD triazine structures in Å

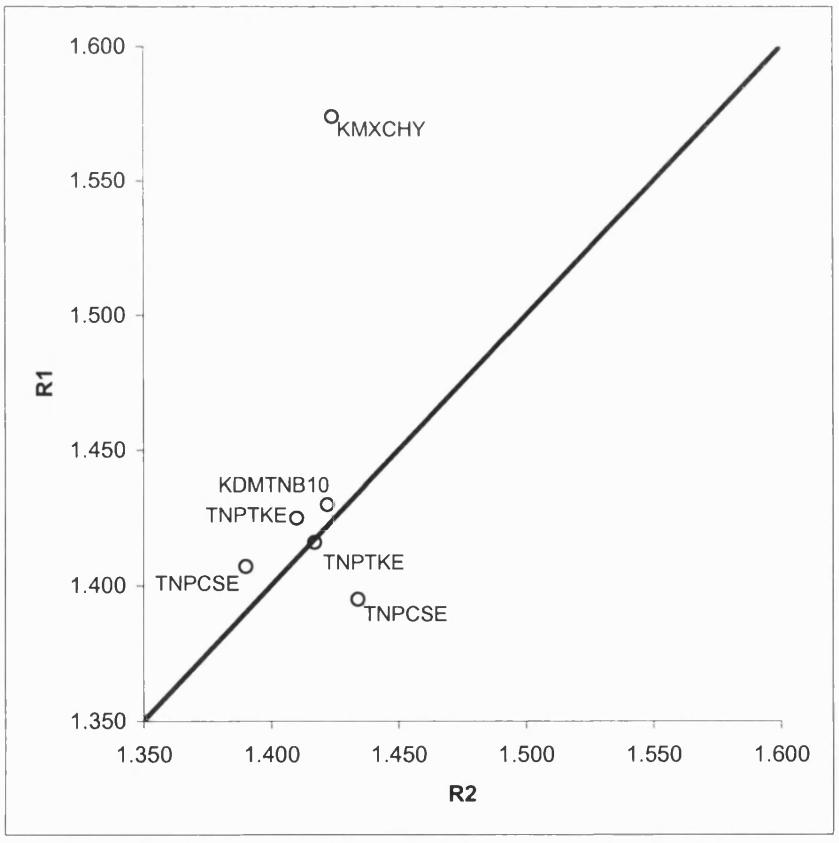


Figure 8.7 – Plot of R1 versus R2 for the CSD TNB structures in Å

8.4) Computational methodology

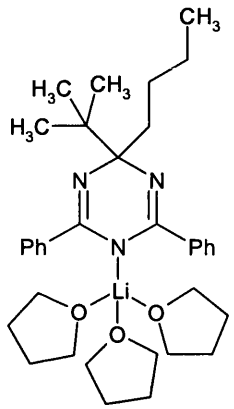
The structures presented here were geometry-optimised. All the structures (1,1'-dimethyl-1,3,5-triazine, 1,1'-dimethoxy-1,3,5-triazine and 1,1'-dimethoxy-2,4,6-trinitrobenzene) were specified using Z-matrix coordinates and symmetry was used wherever appropriate during optimisation. The calculations were performed using the DFT methods employed within the Gaussian 98 package[46]. The following combinations of density functionals and basis sets were used; BLYP/6-31+G(d), BLYP/6-311++G(d,p), B3LYP/6-31+G(d) and B3LYP/6-311++G(d,p). Additionally the BLYP/6-31+G and BLYP/6-31G(d) methods were used for 1,1'-dimethyl-1,3,5-triazine. The opt=Tight keyword was specified initially for all the systems, significantly increasing optimisations convergence criteria significantly reducing the maximum force, RMS force, maximum displacement and RMS displacement required to cease optimising. When it was not possible to optimise this tight requirement was removed especially for the larger TNB systems. The 1,1'-dimethoxy-2,4,6-trinitrobenzene entity was also calculated using the most sophisticated method, B3LYP/6-311++G(d,p), in the presence of solvent using the COSMO Polarised Conductor Model[59] (CPCM) using the parameterisation for solvation in water for comparative purposes.

8.5) Results and discussion

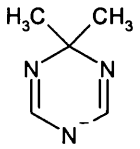
The calculated bond lengths are compared with those for both the most representative (most similar to the calculated structure) CSD structure and to the mean geometry of either triazine or TNB based CSD structures. The most representative structure from the triazine CSD search was LAFBUP10 ((4-^{n}butyl-4-^{t}butyl-2,6-diphenyl-1,4-dihydro-1,3,5-triazine)-tris-(tetrahydrofuran)-lithium.

LAFBUP10 and the mean geometry of the structures of the CSD search are compared here to the calculated gas phase 1,1'-dimethyl-1,3,5-triazine structure, shown in figure 8.8. The bond lengths within the ring (B1, B2 and B3) and the mean bond length from the ring atom to the binding groups (R) for the triazinyl structures are shown in Figure 8.10. In all cases (triazinyl and TNB) the bonds described are as specified in Figure 8.1. The most representative structure for the TNB CSD search was 1,1'-dimethoxy-2,4,6-trinitrobenzene potassium dihydrate (KDMTNB10); The values for this structure and the mean of the values from the

CSD structure are compared to 1,1'-dimethoxy-2,4,6-trinitrobenzene as calculated by gas phase BLYP and B3LYP and aqueous B3LYP DFT approaches. The TNB most representative and TNB calculated structures are shown in figure 8.9. Additionally, the phenyl and triazinyl rings are compared to the calculation of 1,1'-dimethoxy-1,3,5-triazine using gas phase BLYP and B3LYP DFT approaches. The bond lengths within the ring (B1, B2 and B3) and the mean bond length to the binding groups from the ring atom (R) for the calculated and CSD TNB structures are shown in Figure 8.10.

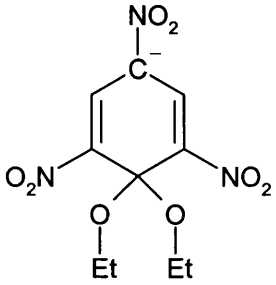


LAFBUP10

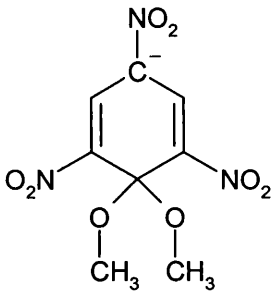


Calculated analogue

Figure 8.8 – The structure of the most representative triazinyl structure LAFBUP10 (((4-^{n}butyl-4-^{t}butyl-2,6-diphenyl-1,4-dihydro-1,3,5-triazine)-tris-(tetrahydrofuran)-lithium) and it calculated analogue 1,1'-dimethyl-1,3,5-triazine.



KDMTNB10



Calculated analogue

Figure 8.9 – The structure of the most representative structure KDMTNB10 (1,1'-dimethoxy-2,4,6-trinitrobenzene potassium dihydrate) and it calculated analogue 1,1'-dimethoxy-2,4,6-trinitrobenzene.

Structure	Symbol	B1	B2	B3	R	A1	A2	A3
CSD (LAFBUP10)		1.467	1.289	1.379	1.547	111.1	108.5	106.3
CSD (Average)		1.476	1.289	1.377	1.548	111.3	107.9	107.3
1,1'-dimethyl-1,3,5-triazine anion								
AM1		1.484	1.320	1.378	1.561	104.7	108.5	108.5
BLYP/6-31+G(d)		1.495	1.318	1.377	1.558	108.9	108.2	108.3
BLYP/6-311++G(d,p)		1.494	1.314	1.375	1.556	109.1	108.2	108.3
B3LYP/6-31+G(d)		1.478	1.304	1.364	1.545	108.9	108.3	108.3
B3LYP/6-311++G(d,p)		1.477	1.300	1.362	1.543	109.0	108.3	108.3

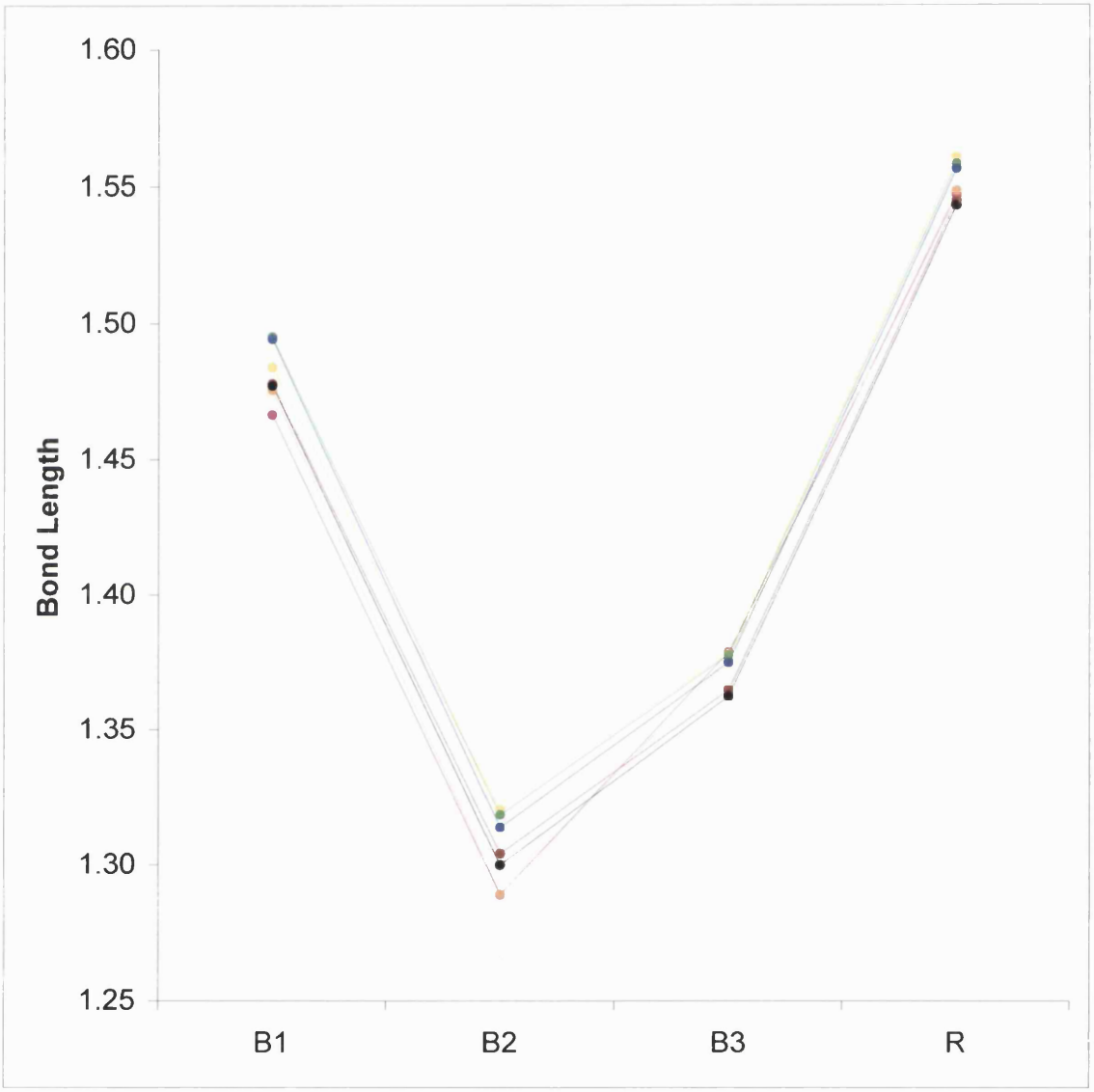


Figure 8.10 - Bond lengths (in Å) for CSD structures and computationally calculated 1,1'-dimethyl-1,3,5-triazine structure. The lines are included for ease of viewing.

The neglect of crystal packing effects and environmental effects limit the possible comparison for both the triazinyl and TNB structures. The CSD structures have high R-factors, especially for the TNB CSD structures which commonly have R-factors of greater than 0.1. However, the comparison of the CSD structures to calculated structures provides one of the only methods to assess how well theoretical methods reproduce real-world geometries.

The triazinyl structures, Figure 8.10, show a close correlation between calculated and CSD bond lengths with the principal features of the CSD being reproduced by the calculated. As expected, the bonds from the substituents binding to the ring, the mean value R is given is the longest bond in both the calculated and crystal structures. R1 and R2 have similar although not identical bond lengths; this difference is due to a buckling in the ring connection to the tetravalent carbon, by approximately 11°, the DFT calculations reproduce this effect. Experimentally this effect has been found in ring systems, for example in the pyridine substituents in the crystal structure of 2,4,5-tris(2-pyridyl)-1,3,5-triazine there was a distinct distortion from planarity shown with the least–squares mean planes of the pyridine rings twisted from the plane of the central triazine by 15.7, 33.8 and 19.8 degrees. [164]The AM1 calculations, however calculated the ring the to be planar, with R1 equal to R2, which is incorrect compared with the experimental structures. This inability to predict the triazinyl ring structure is the reason the TNB structures were only studied using DFT approaches. The mean bond length to the leaving group, R, was experimentally 1.547Å and 1.548Å in LAFBUP10 and average CSD structures respectively; this was consistently overestimated by the computational approaches ranging from 1.543Å to 1.561Å. The B3LYP DFT approach was closer to experiment than the BLYP approach.

The ring bond lengths were paired due to the symmetry present in the calculated molecule in the gas phase calculations with B1 equal to B6, B2 equal to B5 and B3 equal to B4. Therefore only B1, B2 and B3 bond lengths are presented. These were close but not identical in the CSD structures due to crystal packing interactions and as they did not exhibit perfect symmetry. The calculated structures reproduced the experimental trend where B1 to the sp^3 hybridised carbon were the longest, B3 were substantially shorter (by ~0.1Å) and B2 were the shortest (by a further ~0.05Å). The calculated B1 and B2 bond lengths are longer than the CSD structures, whereas B3 was similar to the experimental values. The calculated structures overestimated the bond lengths with the most intensive approach, B3LYP density functional, producing the closest bond lengths to the crystal structures.

Structure	Symbol	B1	B2	B3	R	A1	A2	A3
CSD (KDMTNB10)		1.496	1.356	1.394	1.426	100.1	112.9	114.5
CSD (Average)		1.505	1.356	1.404	1.429	100.6	112.3	112.3
1,1'-dimethoxy-2,4,6-trinitrobenzene anion								
BLYP/6-31+G(d)		1.543	1.377	1.414	1.442	100.7	115.1	108.9
BLYP/6-311++G(d,p)		1.543	1.377	1.414	1.442	100.7	115.2	108.8
B3LYP/6-31+G(d)		1.535	1.370	1.408	1.421	101.4	114.3	109.4
B3LYP/6-311++G(d,p)		1.533	1.365	1.405	1.420	101.4	114.6	109.2
B3LYP/6-311++G(d,p)/CPCM		1.533	1.355	1.447	1.443	104.9	103.7	111.4

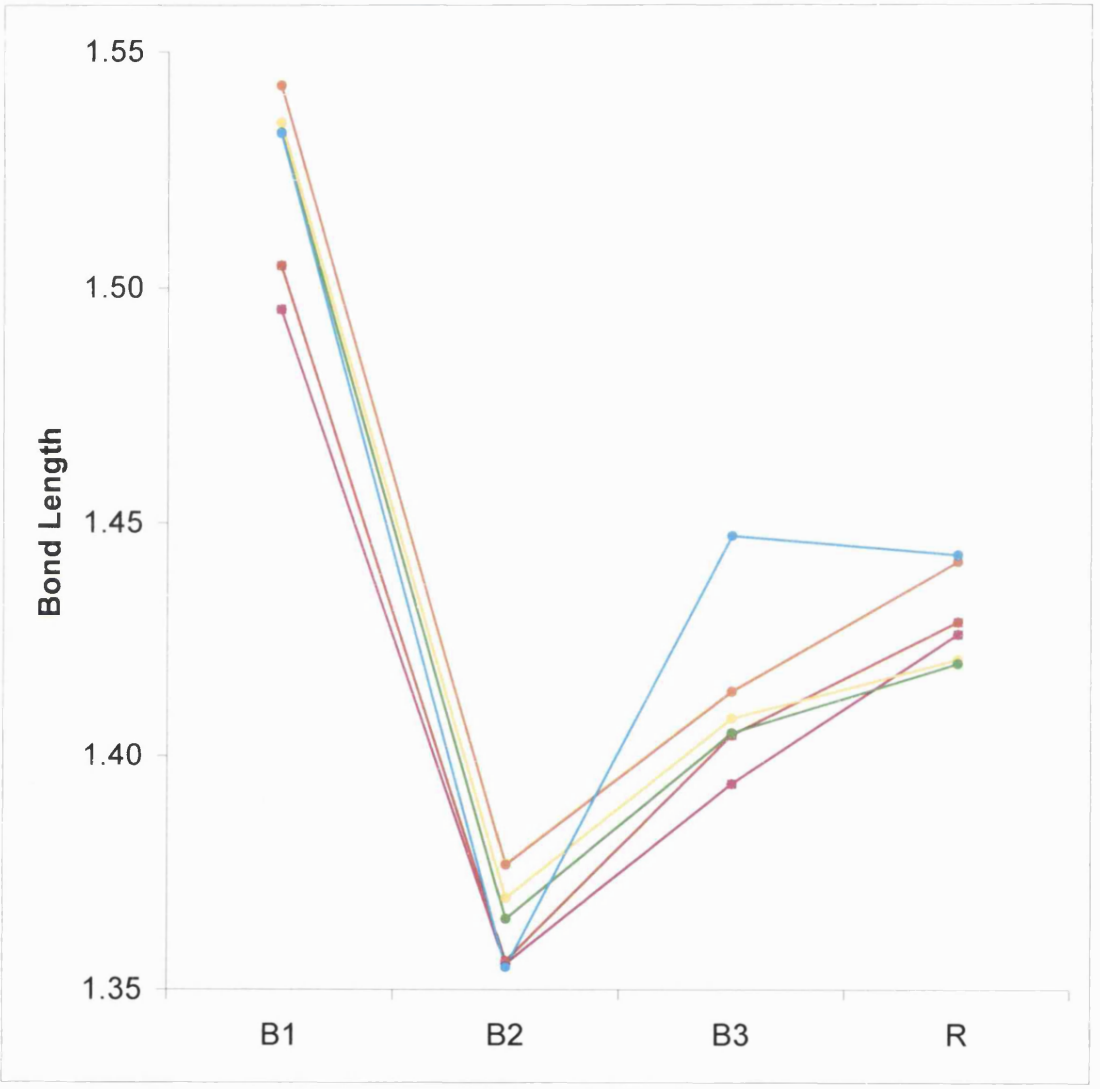


Figure 8.11 – Bond lengths (in Å) for CSD structures, computationally calculated 1,1'-dimethoxy-2,4,6-trinitrobenzene structures. The lines are included for ease of viewing.

Structure	Symbol	B1	B2	B3	R	A1	A2	A3
CSD (KDMTNB10)	■	1.496	1.356	1.394	1.426	100.1	112.9	114.5
CSD (Average)	■	1.505	1.356	1.404	1.429	100.6	112.3	112.3
1,1'-dimethoxy-1,3,5-triazine anion								
BLYP/6-31+G(d)	●	1.452	1.324	1.369	1.466	98.0	109.7	109.8
BLYP/6-311++G(d,p)	●	1.450	1.320	1.366	1.466	98.1	109.8	109.8
B3LYP/6-31+G(d)	●	1.441	1.309	1.356	1.439	98.9	109.8	109.8
B3LYP/6-311++G(d,p)	●	1.440	1.305	1.353	1.439	99.0	109.8	109.8

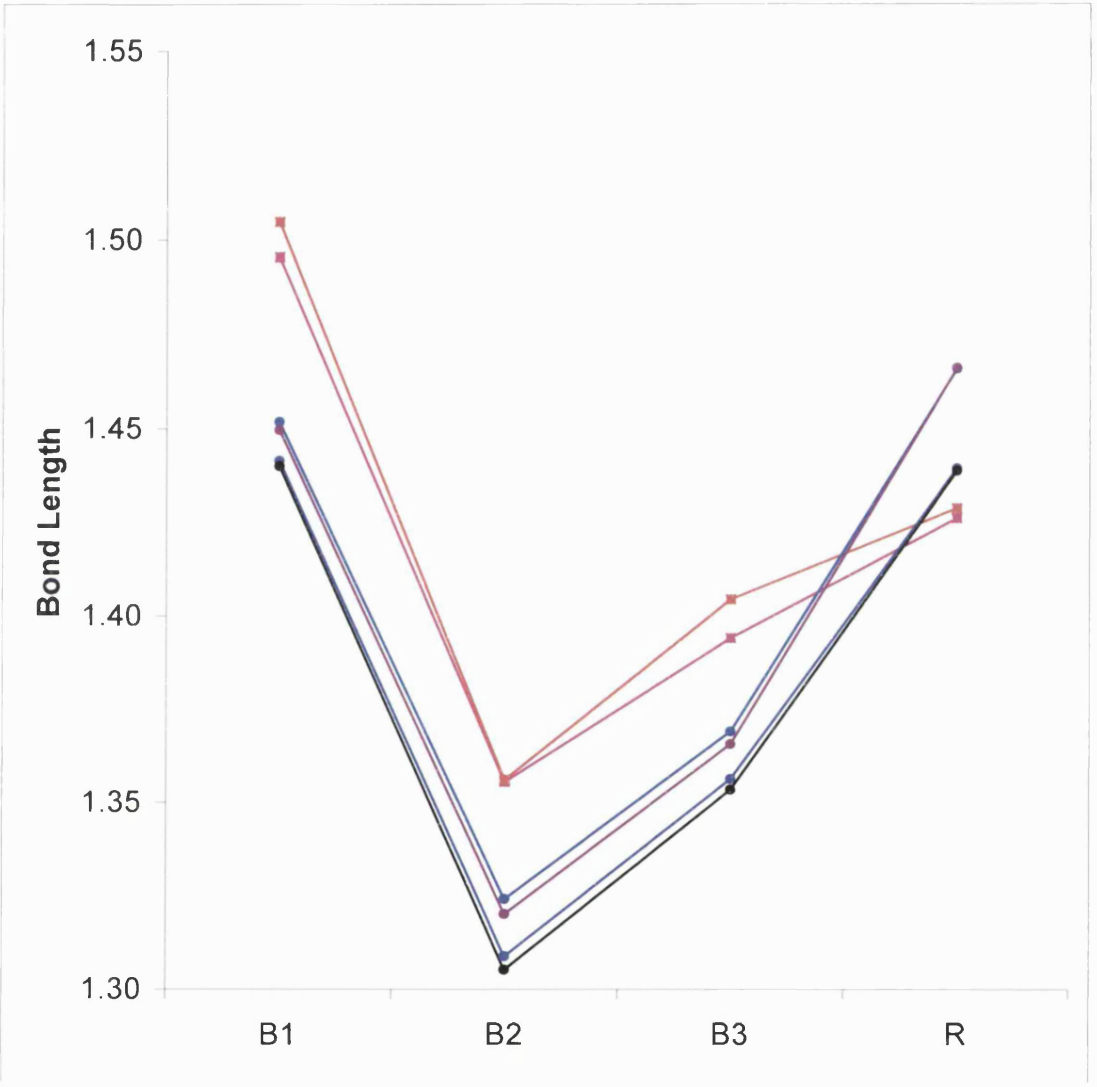


Figure 8.12 – Bond lengths (in Å) for CSD structures, computationally calculated 1,1'-dimethoxy-1,3,5-triazine structures. The lines are included for ease of viewing.

The TNB structures, Figure 8.11 and 8.12, like the triazinyl structures, show a close correlation between calculated and crystallographic bond lengths, even with the neglect of environment effects; with the structure calculated here (1,1'-dimethoxy-2,4,6-trinitrobenzene) close to the most representative CSD structure 1,1'-dimethoxy-2,4,6-trinitrobenzene potassium dihydrate (KDMTNB10).

The TNB bond lengths for atoms within the ring are paired due to symmetry B1 equal to B6, B2 equal to B5, B3 equal to B4 and R1 equal to R2; these pairs had similar but non-identical bond lengths in the crystal and solvated structures, due to the environment effects. The gas phase 1,1'-dimethoxy-2,4,6-trinitrobenzene structures closely followed the trend in the CSD bond lengths and these decreased in the following order B1, R, B3 then B2. The aqueous calculation showed B3 to be longer than R and showed uneven bond lengths within the 'pairs' due to intermolecular hydrogen bonding between the hydrogens on the methoxy- and oxygens on the nitro-substituents.

The gas phase 1,1'-dimethoxy-2,4,6-trinitrobenzene calculations followed the trend of the crystal bond lengths well, with the B3LYP/6-311++G(d,p) method best reproducing the crystal geometries. Overall the bond lengths for the B3LYP structures were ~0.01Å shorter than from BLYP. All the ring bond lengths (B1, B2 and B3) are calculated to be longer than in the average crystal structures. The structures calculated using the larger basis set were also closer to the crystal bond lengths, although the effect of the basis set was marginal and much less than that of the change in density functional.

The ring bond lengths, the calculated triazinyl equivalent structure (1,1'-dimethoxy-1,3,5-triazine) were shorter than for the calculated TNB structure (1,1'-dimethoxy-2,4,6-trinitrobenzene). The bond to the nucleophile and leaving group, R, was generally longer in the calculated triazinyl structure indicating weaker bonds and suggesting this intermediate structure is less stable than the TNB structure.

Throughout the series of calculated structures, the B3LYP produced shorter bond lengths compared with the BLYP method and the 6-311++G(d,p) basis set produced shorter and more accurate bond lengths than 6-31+G(d). This effect caused the more computationally expensive B3LYP/6-311++G(d,p) method to be the closest to the experimental structures, for both the triazinyl and TNB

structures. The addition of a solvation model assists in the simulation of the environment effect, however the dielectric used in the CPCM calculation, that of pure water, may be incorrect compared to experimental conditions and thus overestimates this effect.

A range of papers have been published involving the calculation of molecular geometry, for example, the density functional study of an amide. The published bond lengths[165] differ by 0.5-1.7% between gas phase calculations and gas phase methlyamide and the bond lengths presented here by 0.0-2.5% between gas phase calculations and average CSD structures. The experimental bond lengths and percentage error from experiment is shown in Table 8.5. These techniques therefore can be used to calculate bond lengths reliably and do not differ extensively from similar published results

Table 8.5 – Percentage error in between the BLYP/6-31+G(d), BLYP/6-311++G(d,p), B3LYP/6-31+G(d) and B3LYP/6-311++G(d,p) calculated bond lengths and the experimental C-N and C=O methlyamide bond lengths[165] (in Å) and B1/B2/B3 bond lengths in the average CSD 1,1'-dimethoxy-2,4,6-trinitrobenzene and 1,1'-dimethoxy-1,3,5-triazine structures.

Bond	Experiment or Average CSD Bond Length	Computational method			
		BLYP		B3LYP	
		6-31+G(d)	6-311++G(d,p)	6-31+G(d)	6-311++G(d,p)
C-N	1.368	0.48%	0.37%	0.46%	0.52%
C=O	1.212	1.68%	1.04%	0.60%	0.02%
1,1'-dimethoxy-1,3,5-triazine					
B1	1.476	1.29%	1.22%	0.14%	0.07%
B2	1.289	2.25%	1.94%	1.16%	0.85%
B3	1.377	0.00%	0.15%	0.95%	1.10%
1,1'-dimethoxy-2,4,6-trinitrobenzene					
B1	1.505	2.52%	2.52%	1.99%	1.86%
B2	1.356	1.55%	1.55%	1.03%	0.66%
B3	1.404	0.71%	0.28%	0.07%	3.06%

8.6) Conclusion

The DFT calculations reproduce the trends in the crystal structure bond lengths closely. Computationally less intensive AM1 calculations do not represent the ring adequately in this study. The computational methods generally produce longer bonds than are observed in the crystal structures. The gas phase calculations produced bond lengths on average 0.006Å and 0.014Å longer than the mean triazine and TNB crystal bond lengths respectively. The gas phase B3LYP/6-

311++G(d,p) method was the closest to experiment within 0.002Å and 0.007 Å of the mean triazine and TNB crystal bond lengths. The least intensive DFT method BLYP/6-31+G(d) produces the least accurate bond lengths, although this is still good at within 0.015Å and 0.020Å of the mean triazine and TNB bond lengths respectively. The inclusion of a continuum solvent model representing the environment improves the description of the environmental effects; this provides additional problems concerning the permittivity used for the solvent.

8.7) Evaluation

Density functional methods are able to reproduce the structural trends present in the triazinyl and trinitrobenzene crystal structures from the Cambridge Structural Database (CSD). Therefore, the calculated structures produced for Meisenheimer-type compounds with these methods will be a good approximation to experimentally undeterminable structures.

Chapter 9

Modelling triazine and trinitrobenzene derivatives reactivity using density functional methods

Can density functional methods reproduce experimentally derived mechanisms and produce reliable vibrational frequencies?

9.1) Introduction

The fixation mechanism of reactive dyes based on the 1,3,5-triazine unit, in particular the competition between fixation and hydrolysis, is examined using density functional theory (DFT). The conditions that promote fixation over hydrolysis are of industrial interest, for example the effect of leaving groups and solvent on the reaction mechanism. Computational[85-89] and experimental[77-84] studies of nucleophilic substitution and experimental mechanistic studies of the 1,3,5-triazine unit[90-94] have been alluded to previously in Chapters 4 and 5, therefore only a brief outline will be given here. Studies involving nitroarenes are more numerous; these reactions have been shown to proceed via the formation of a σ-adduct - Meisenheimer complex. [7]

It is of particular interest whether fixation proceeds via a stepwise (A_N + D_N Figure 9.1) or a concerted (A_ND_N Figure 9.2) process. Hydrolysis, methanolysis, fluorolysis and chlorolysis of fluorotriazine and chlorotriazine are assessed using DFT, with and without the presence of the CPCM solvent representation.

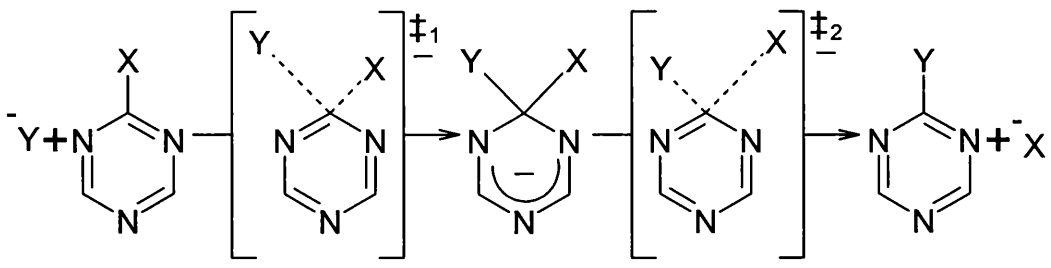


Figure 9.1 – The stepwise (A_N + D_N) mechanism for the nucleophilic attack of monosubstituted triazine ring (X-triazine) by a nucleophile Y.

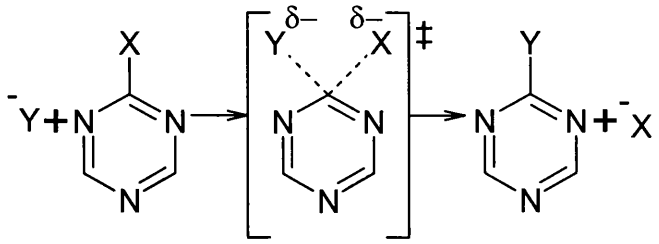


Figure 9.2 – The concerted (A_ND_N) mechanism for the nucleophilic attack of monosubstituted triazine ring (X-triazine) by a nucleophile Y.

Vibrational frequency calculations are performed on chemically significant structures in order to characterise whether the structure is a transition structure or an intermediate. The significance of the presence of a number of imaginary vibrational frequencies (as described in Section 3.2) is that an energy minimum species should have none; a structure with a single imaginary frequency (first order saddle point) is a transition structure, whereas a structure with two imaginary frequencies is a local maximum (second order saddle point). It was found during attempts at characterising the triazinyl species that the PCM and CPCM solvation models produced erroneous results. Therefore, the reliability of frequency calculations using these methods is assessed for a range of small organic compounds.

9.2) Computational methodology

DFT approaches, BLYP/6-31+G(d) and BLYP/6-31+G(d)/CPCM(water), have been utilised to examine the potential energy surfaces (PES) for reactions involving the 1,3,5-triazine moiety. The monosubstituted reactive unit has been focused upon due to the high computational expense of treating the whole reactive dye using DFT methods. The substituents present in the entire dye structure are assumed to perturb the electronic structure and PES produced only slightly. In common with many experimental studies, the cellulose lattice (anionic glucosyl moiety) involved in dye fixation is represented by a methoxide anion; this considerably reduces the complexity of the calculations and computational expense involved.

The DFT method was used as it has been shown to reproduce structures and energetics in a vacuum (chapter 6 and 8). The BLYP functional was employed with the moderate 6-31+G(d) basis set in order to produce computationally

efficient results. It has been shown that there is not a significant difference in the energetics and structures produced by B3LYP or BLYP or with larger basis sets (6-311++G(d,p)), although considerable variation in computational expense is involved.

The DFT calculations were carried out using the Gaussian 98[46] computational package. This yields aqueous Gibbs free energy of solvation, ΔG_{solv}. These methods were employed along with full geometry optimisations to obtain the energetics of the reactant, product and 'intermediate' species. The reaction paths from the intermediate to both the reactant and product species were followed for the identity reactions by gradually increasing the bond length to the leaving group. It was, however, not possible to find any further stationary points.

Characterisation of stationary points through the calculation of vibrational frequencies was not possible for the aqueous species due to the production of anomalous results. In view of this, the frequency calculations were then investigated via changing the stepsize in the numerical differentiation used for a variety of simple compounds in the gas phase and using both the CPCM and PCM continuum solvation models.

9.3) Results and discussion

All energy differences stated here are in kilojoules per mole and all total energies are given in atomic units (1 kJ mol^{-1} = 0.239 kcal mol^{-1} = $3.81x10^{-4}$ Hartrees). All the frequencies calculated are shown in wavenumbers (cm^{-1}) with imaginary frequencies shown as negative values; where the number produced was too large to be displayed in the output, asterisks are shown. All the reactions were calculated using the BLYP/6-31+G(d), and aqueous phase, BLYP/6-31+G(d)/CPCM(water) DFT methods within Gaussian 98. The total energies, in Hartrees, for each of the individual species (reactant, product and intermediate) are shown in Table 9.1. The total energies, in Hartrees, of the combined reactants and products along with the intermediates for all the reactions as well as the differences, in kJ mol^{-1}, between the intermediate and the reactants and the products and the intermediates are shown in Appendix B.

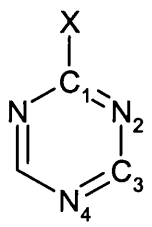
Analytical, numeric and default frequencies for formaldehyde using the HF/6-31G(d), HF/6-31+G(d,p), B3LYP/6-31G(d) and B3LYP/6-31+G(d,p) methods are shown in Appendix C. The BLYP/6-31+G(d), BLYP/6-31+G(d)/CPCM(water), BLYP/6-31+G(d)/PCM(water) frequencies and total energies for a series of small organic species are shown in Appendix D.

Table 9.1 –The total energies, in Hartrees, for the reactant and intermediate (σ-adduct or ion-molecule complex) structures both in the gas and aqueous phases.

	Computational method	
	BLYP/6-31+G(d)	**BLYP/6-31+G(d)/CPCM(water)**
Reactant		
F^-	-99.8539	-100.0202
Cl^-	-460.2528	-460.3723
^-OH	-75.7859	-75.9617
^-OMe	-115.0746	-115.2143
tr_F	-379.5509	-379.5642
tr_Cl	-739.8848	-739.8969
tr_OH	-355.5337	-355.5541
tr_OMe	-394.8117	-394.8255
Intermediate (σ-adduct or ion-molecule complex) ‡		
F tr $F^{-\#}$	-479.4642	-479.5599
Cl tr $Cl^{-\#}$	-1200.1358	-1200.2238
OH tr $OH^{-\#}$	-431.4914	-431.4910
F tr $OH^{-\#}$	-455.4270	-455.5261
F tr $OMe^{-\#}$	-494.6986	-494.7995
F tr Cl^{-*}	-839.8134	-839.9194
OH tr Cl^{-*}	-815.8245	-815.9099
OMe tr Cl^{-*}	-855.0798	-855.1852

‡ The intermediates shown are σ-Adduct$^{\#}$ for the identity reactions and a Cl^-.....Tr-Y^*

Table 9.2 –The Mulliken[166-169] charges with the hydrogen charge summed on the heavy atoms chlorotriazine and fluorotriazine in both the gas and aqueous phases.

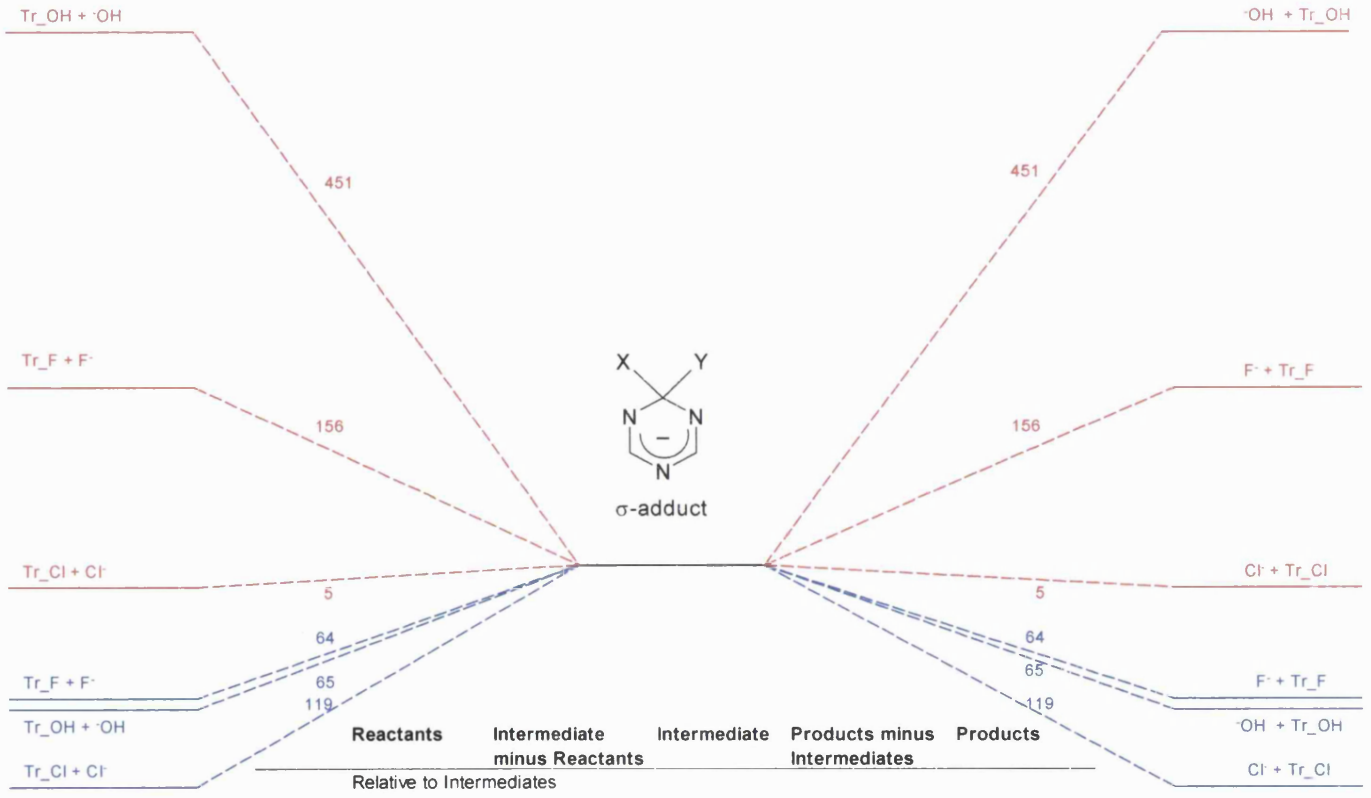


Computational Method	**Halogen**	**Mulliken charge on heavy atoms (incorperating H charge)**				
	X	C_1	N_2	C_3	N_4	X
BLYP/6-31+G(d)	Cl	-0.229	-0.067	0.223	-0.164	0.081
BLYP/6-31+G(d)	F	0.459	-0.216	0.194	-0.174	-0.240
BLYP/6-31+G(d)/CPCM(Water)	Cl	-0.163	-0.127	0.312	-0.233	0.027
BLYP/6-31+G(d)/CPCM(Water)	F	0.520	-0.279	0.278	-0.245	-0.272

9.3.1) The hydrolysis of hydroxytriazine, the fluorolysis of fluorotriazine and the chlorolysis of chlorotriazine

The vacuum and aqueous phase DFT reaction energy profiles for the identity mechanisms (hydrolysis of hydroxytriazine, fluorolysis of fluorotriazine and chlorolysis of chlorotriazine) relative to the adduct structure are shown schematically in Figure 9.3. All the structures, in particular the reactant and product structures, are stabilised in the aqueous phase compared with vacuum calculations. The formation of the σ-adduct requires energy in all the aqueous phase calculations, suggesting that the adduct is a transition structure for the identity nucleophilic substitution mechanisms. In the gas phase, the chlorolysis of chlorotriazine was the only mechanism that required energy to form the adduct, with both the fluorolysis and hydrolysis releasing considerable energy in this step.

The relative stability of the reactants (and products) in the gas phase follow the same trend as the gas phase basicity of the attacking nucleophile. The gas phase basicity for hydroxide ($^{-}$OH), fluoride (F^{-}) and chloride (Cl^{-}) were -1607, -1530 and -1373 kJ mol^{-1}[170] respectively. The most basic nucleophile releases the most energy upon formation of the adduct showing a energy change of -451, -156 and 5 kJ mol^{-1} for hydrolysis ($^{-}$OH), fluorolysis (F^{-}) and chlorolysis (Cl^{-}) respectively for the vacuum calculations. The effect of solvation stabilises all the structures, in particular the hydrolysis reactants (and products). The hydroxide anion, was stabilised significantly by the continuum compared to the respective adduct. This causes the formation of adducts by hydrolysis and fluorolysis to have similar energy differences requiring 65 and 64 kJ mol^{-1} respectively, whereas the chlorolysis requires significantly more energy, 119 kJ mol^{-1}. All the adducts for these identity reactions are σ-adducts with covalent like bond lengths to the binding and leaving group (none of these formed adducts resembling ion-molecule complexes).



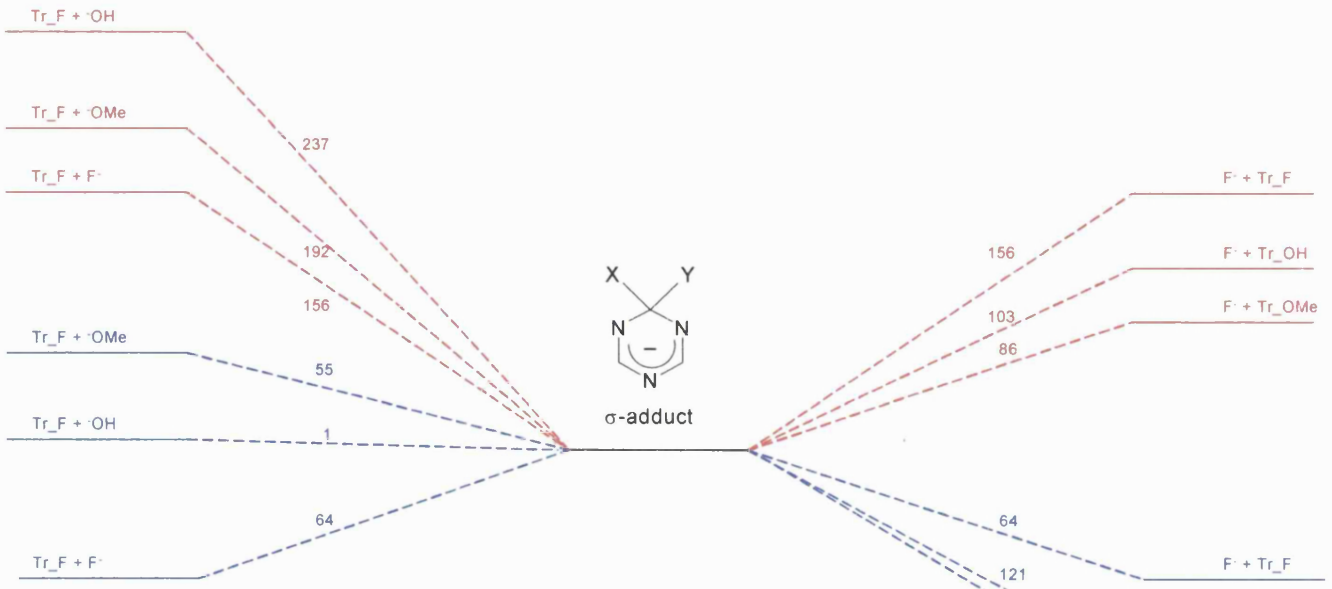
	Reactants		Intermediate minus Reactants		Intermediate	Products minus Intermediates		Products
	Relative to Intermediates							
Vacuum	tr_F	F⁻	-155.7	F tr F⁻		155.7	F⁻	tr_F
Vacuum	tr_Cl	Cl⁻	4.8	Cl tr Cl⁻		-4.8	Cl⁻	tr_Cl
Vacuum	tr_OH	⁻OH	-451.1	OH tr OH⁻		451.1	⁻OH	tr_OH
Aqueous	tr_F	F⁻	64.2	F tr F⁻		-64.2	F⁻	tr_F
Aqueous	tr_Cl	Cl⁻	119.1	Cl tr Cl⁻		-119.1	Cl⁻	tr_Cl
Aqueous	tr_OH	⁻OH	64.9	OH tr OH⁻		-64.9	⁻OH	tr_OH

Figure 9.3 – The vacuum, BLYP/6-31+G(d), and aqueous phase, BLYP/6-31+G(d)/CPCM(water), reaction energy profiles for the identity nucleophilic substitution of monosubstituted-1,3,5-triazine (X-triazine) by a nucleophile Y.

9.3.2) The methanolysis, hydrolysis and fluorolysis of fluorotriazine

The vacuum and aqueous phase DFT reaction energy profiles relative to the adduct structure for the methanolysis, hydrolysis and fluorolysis of fluorotriazine are shown schematically in Figure 9.4. All the structures are stabilised by the presence of solvent, compared to the vacuum, with the reactant and product species stabilised to a greater degree than the adduct structure. In all the systems, the reactants are equal or higher in energy than the product structures.

The relative stability of the reactants follow the same trend as the gas basicity of the nucleophile, as in the identity reactions. The gas basicity for hydroxide (^-OH), methoxide ($^-OCH_3$) and fluoride (F^-) were -1607, -1559 and -1530 kJ mol^{-1}[170] respectively. This shows that the most basic nucleophile is most ready to combine to form an adduct complex, as in the identity reactions. The gas phase energy change upon the formation of the adducts is –237, -192, -156 kJ mol^{-1} for the hydrolysis (^-OH), methanolysis ($^-OCH_3$) and fluorolysis (F^-) of fluorotriazine respectively. Whereas in the aqueous phase, the formation of the methanolysis adduct, rather than the hydrolysis, releases the most energy with the energy change upon formation being –55, -1 and 64 kJ mol^{-1} for methanolysis ($^-OCH_3$), hydrolysis (^-OH) and fluorolysis (F^-) of fluorotriazine respectively. The extra stabilisation of the hydrolysis compared to the methanolysis reactants relative to the adduct species, is due to the hydroxide anion being stabilised more by the continuum than the more diffuse methoxide anion. The methanolysis reaction has the greatest release of energy of reaction upon formation of the product from the adduct structure, in both vacuum and aqueous phase. The formation of the adduct and the formation of the products from the adduct for the fluorolysis (identity) reaction was least favourable in both the gas and aqueous phase. The semiempirical AM1 and AM1/SM2.1 studies of the hydrolysis and methanolysis of fluorotriazine, shown in Figures 4.4 and 4.6 and described in chapter 4. These show a larger energy difference for the formation of the intermediate and from the formation of the products from the intermediate the DFT methods.



	Reactants		Intermediate minus Reactants	Intermediate	Products minus Intermediates	Products
	Relative to Intermediates					
Vacuum	tr_F	F⁻	-155.7 F tr F⁻		155.7 F⁻	tr_F
Vacuum	tr_F	⁻OH	-236.8 F tr OH⁻		103.4 F⁻	tr_OH
Vacuum	tr_F	⁻OMe	-191.8 F tr OMe⁻		86.5 F⁻	tr_OMe
Aqueous	tr_F	F⁻	64.2 F tr F⁻		-64.2 F⁻	tr_F
Aqueous	tr_F	⁻OH	-0.6 F tr OH⁻		-126.4 F⁻	tr_OH
Aqueous	tr_F	⁻OMe	-55.2 F tr OMe⁻		-121.2 F⁻	tr_OMe

Figure 9.4 – The vacuum, BLYP/6-31+G(d), and aqueous phase, BLYP/6-31+G(d)/CPCM(water), reaction energy profiles for the hydrolysis, methanolysis and fluorolysis of monofluoro-1,3,5-triazine.

9.3.3) The methanolysis, hydrolysis, fluorolysis and chlorolysis of chlorotriazine

The vacuum and aqueous phase DFT reaction energy profiles relative to the *reactants* for the methanolysis, hydrolysis, fluorolysis and chlorolysis of chlorotriazine are shown schematically in Figure 9.5. Throughout these reactions, the intermediate energy is lower relative to the reactants in the vacuum than in the aqueous phase. All the non-identity reactions (hydrolysis, methanolysis and fluorolysis) formed chloride ion-molecule complexes, rather than σ-adducts. No transition structures between the chloride ion-molecule complexes and the reactant and product structures were found. Some stationary points have to be present along the reaction coordinate in the aqueous phase, although these could have a small energy difference. In vacuum, the nucleophile could just be subsumed in a barrierless reaction energy process.

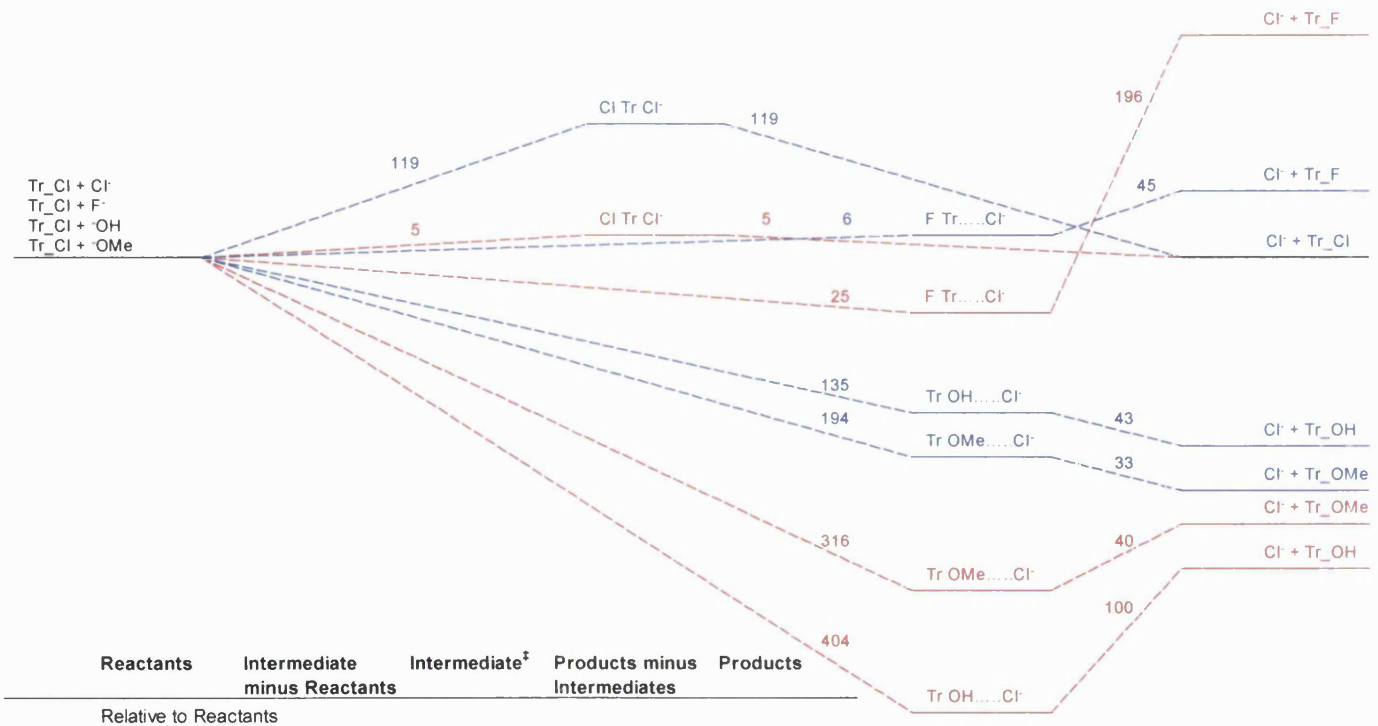
The chloride ion-molecule complexes show the chloride anion interacting electrostatically with either the chlorine of the triazinyl moiety or with a substituent group. These charge-dipole interactions to the substituent, hydroxide or methoxide group, provide additional stabilisation for the respective complexes in vacuum.

The relative stability of the reactants follows the same trends as in the previous Section. The stability of the gas phase reactants correlated strongly with the gas basicity of the nucleophile. The gas phase hydrolysis of chlorotriazine releases more energy than methanolysis, whereas the opposite preference is shown in the aqueous calculations. The gas phase energy change upon formation of the intermediate for the hydrolysis and methanolysis of chlorotriazine showed energy differences of -404 kJ mol^{-1} and -316 kJ mol^{-1} respectively; whereas in the aqueous phase formation of the intermediate for the hydrolysis and methanolysis of chlorotriazine gave values of –135kJ mol^{-1} and -194 kJ mol^{-1} respectively. The change of preference for this step could be due to the higher concentration of charge on hydroxytriazine than methoxytriazine anion. This leads to a stronger ion-molecule bond in hydroxytriazine than methoxytriazine, especially in the gas phase. Both the structures are stabilised by the presence of solvent, with both energies being closer than in the gas phase. The products of the hydrolysis and

methanolysis of chlorotriazine for the gas phase are substantially higher in energy than the chloride ion-molecule complex and lower in the aqueous calculations.

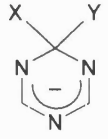
The reaction energy profile for the fluorolysis of chlorotriazine shows marked differences to the hydrolysis and methanolysis reactions. The aqueous phase chloride ion-molecule complex is higher in energy than the reactants by 6 kJ mol^{-1} and only 25 kJ mol^{-1} lower in energy in the gas phase. Both the gas and aqueous phase products are significantly higher in energy than the ion-molecule structures and the reactants. The energy to change from the ion-molecule complex to products for the gas and aqueous calculations was 196 and 45 kJ mol^{-1} respectively.

The products of the fluorolysis of chlorotriazine (fluorotriazine and chloride ion) are significantly higher in energy (relative to the reactants) than all the other reactions, including the products of the chlorolysis identity reaction (chlorotriazine and chloride ion). The only difference between these products is the difference in energy between the chlorotriazine and fluorotriazine species. This energy difference is due to the distribution of charge within these species. The Mulliken[166-169] charges on the heavy atoms (the charge on the hydrogens are summed onto the heavy atoms) in both the gas and aqueous phases are shown in Table 9.2. These charges clearly show that the chlorine is less electron withdrawing than fluorine, as expected from the relative electronegativities. The charge on the fluorotriazine therefore has a higher variation between the alternating carbons and nitrogens, higher nodality, than chlorotriazine. A higher nodality decreases the stability of a compound triazine; for example, is destabilised in comparison to benzene. This effect destabilises and increases the energy of fluorotriazine relative to chlorotriazine in both the gas and aqueous phases. The effect of solvation increases the nodality in both rings, however both the charges on fluorotriazine and chlorotriazine are stabilised by the continuum. This reduces the differences between the charge distribution and energetics of the two structures. There is a pronounced change of charge on C_1 for both the structures; it is strongly positive for fluorotriazine and slightly negative for chlorotriazine. This may indicate why the different intermediate structures are formed as the chloride ion forms an ion-molecule complex with fluorotriazine binding to the negative C_1 and a σ-adduct for chlorolysis of fluorotriazine where this is not possible.

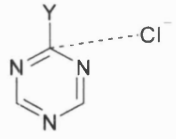


	Reactants		Intermediate minus Reactants	Intermediate‡	Products minus Intermediates	Products
	Relative to Reactants					
Vacuum	tr_Cl	Cl⁻	4.8	Cl tr Cl⁻#	-4.8 Cl⁻	tr_Cl
Vacuum	tr_Cl	F⁻	-25.3	F tr...Cl⁻*	196.1 Cl⁻	tr_Cl
Vacuum	tr_Cl	⁻OH	-403.9	OH tr...Cl⁻*	99.8 Cl⁻	tr_OH
Vacuum	tr_Cl	⁻OMe	-316.2	OMe tr...Cl⁻*	40.2 Cl⁻	tr_OMe
Aqueous	tr_Cl	Cl⁻	119.1	Cl tr Cl⁻#	-119.1 Cl⁻	tr_Cl
Aqueous	tr_Cl	F⁻	6.0	F tr...Cl⁻*	44.9 Cl⁻	tr_F
Aqueous	tr_Cl	⁻OH	-134.6	OH tr...F⁻*	-43.2 Cl⁻	tr_OH
Aqueous	tr_Cl	⁻OMe	-194.3	OMe tr...Cl⁻*	-33.0 Cl⁻	tr_OMe

‡ The intermediates shown are σ-Adduct# and choride ion-molecule complexes (Cl⁻.....Tr-Y*)



σ-adduct



Chloride ion-molecule complex

Figure 9.5 – The vacuum, BLYP/6-31+G(d), and aqueous phase, BLYP/6-31+G(d)/CPCM(water), reaction energy profiles for the hydrolysis, methanolysis, chlorolysis and fluorolysis of monochloro-1,3,5-triazine.

9.3.4) Vibrational frequencies

It was found during attempted characterisation of the adduct species from the triazinyl reactions (Sections 9.3.1, 9.3.2 and 9.3.3), that CPCM implemented within Gaussian 98 clearly produced incorrect results in the calculation of vibrational frequencies by numerical differencing (the only option available for PCM and CPCM solvation models in Gaussian 98). In these triazinyl molecules, a vibrational out of plane mode couples with another similar mode giving a pair of frequencies with one very high and one very low (and often imaginary). This casts considerable doubt on the reliability of characterisation of transition structures and intermediates in these reactions with this method. The effect of changing the stepsize and therefore the coarseness of the numerical differentiation was assessed.

The vibrational frequencies for gas phase numerical and analytical HF/6-31G(d), HF/6-311+G(d,p), B3LYP/6-31G(d) and B3LYP/6-311+G(d,p) computational methods for formaldehyde is shown in Appendix C. These show minor differences in the frequencies for these gas phase calculations. The total energy, solvated free energy, energy difference and vibrational frequencies for a series of small organic compounds are shown in Appendix D. The formyl anion, formyl fluoride, protonated formamide cation, chloride ion-methyl chloride complex, triazine and the two σ-adduct triazinyl anions for the chlorolysis of chlorotriazine and fluorolysis of fluorotriazine reactions (shown in Section 3.3.1) are the species shown in Appendix D. All the frequencies showed little or no change when different step sizes were used for the gas phase calculations, so only a stepsize of 10 (the default) is shown. This verifies the validity of frequency calculations in the gas phase.

The frequencies for hydrogen cyanide showed very little difference when the stepsize was changed in any of the methods used. The calculated vibrational frequencies were also close to experiment[171] for all the methods used, with the default stepsize with the gas phase considerably closer than the solvated approaches. The calculated and vibrational frequencies for hydrogen cyanide are shown in Table 9.3.

Table 9.3 –The calculated and experimental[171] vibrational frequencies for the π -bend, CN stretch and CH stretch modes in hydrogen cyanide, in cm^{-1}.

Computational Method	HCN π-bend	CN stretch	CH stretch
Experimental	712	2097	3311
BLYP/6-31+G(d)	705	2110	3376
BLYP/6-31+G(d)/PCM(Water)	723	2111	3295
BLYP/6-31+G(d)/CPCM(Water)	726	2109	3290

The frequencies for the remaining compounds showed considerable errors, especially with the shortest and longest stepsize calculated (1 and 100) and for the charged systems. The most prominent errors were also present in the larger triazinyl systems, which are of the most interest here. In a blind test, two near identical structures for the adduct from the nucleophilic attack of chlorotriazine with chloride (Cl tr Cl^-) produced markedly different results, regardless of stepsize. Varying the stepsize for this structure produced erroneous results, where between 2 and 5 imaginary frequencies were shown for the same structure - as shown in Table 9.4. The variation of the stepsize did not manage to produce more realistic values for the imaginary frequencies.

Table 9.3 – The calculated vibrational frequencies for BLYP/6-31+G(d) and BLYP/6-31+G(d)/CPCM(water) for the lowest 6 of 24 vibrational modes of the adduct of the attack of chlorotriazine by a chloride anion, in cm^{-1}. The imaginary frequencies are shown in bold.

Triazine Cl Cl^- Stepsize	Computational Method	Frequencies 1	2	3	4	5	6=>24
10	BLYP/6-31+G(d)	**-150**	123	193	194	259	298
1	BLYP/6-31+G(d)/CPCM(water)	**-3393**	**-2689**	**-1434**	**-1229**	**-380**	189
5	BLYP/6-31+G(d)/CPCM(water)	**-1480**	**-1014**	**-498**	**-162**	191	229
10	BLYP/6-31+G(d)/CPCM(water)	**-682**	**-599**	**-290**	103	182	225
50	BLYP/6-31+G(d)/CPCM(water)	**-7222**	**-4722**	**-2889**	-177	210	277
100	BLYP/6-31+G(d)/CPCM(water)	**-250**	**-240**	47	186	227	296

9.3.5) Path calculations

As it was found that it was not possible to characterise the triazinyl adduct species by using vibrational frequencies, as shown in Section 9.3.4, in order to indicate the nature of the adduct species the structure was perturbed along the reaction coordinate. This was achieved by optimising the structure and then elongating and fixing one of the bonds then re-optimising the structure. A increase in the energy when the structure is perturbed indicates the adduct species is an intermediate. Whereas a decrease in the energy when the structure is perturbed indicates the adduct species is a transition structure.

Table 9.4 – The indication by the reaction profile produced by the elongating of the adducts species bond and subsequent optimisation using BLYP/6-31+G(d) and BLYP/6-31+G(d)/CPCM(water).

Reaction	Vacuum	Aqueous
Fluorolysis of flourotriazine	Stepwise	Stepwise
Chlorolysis of chlorotriazine	Concerted	Stepwise
Hydrolysis of hydroxytriazine	Stepwise	Stepwise
Methanolysis of methanoxytriazine	Stepwise	Stepwise

This indicates all the reactions except for the gas phase chlorolysis of chlorotriazine proceed via a stepwise mechanism. This would indicate that in solution the reactions behave similarly to the benzanoid Meisenheimer type mechanism. These indications cannot be proven absolutely due to the inability to characterise the structures from calculation of vibrational frequencies, however are the closest that we can get given the present implementation of the methodology.

9.4) Conclusion

The effect of solvent on the competition between hydrolysis and methanolysis of monosubstituted triazines is highly significant. The hydrolysis reaction was preferred over methanolysis in the gas phase, whereas methanolysis was preferred over hydrolysis in the aqueous phase. This change in preferences is due to the relative stability of the methoxide and hydroxide anions, which are present in the non-adduct species. The methoxide anions, with more diffuse and delocalised charge and a higher solvent accessible surface area than the

hydroxide anion, are more stabilised owing to the presence of charge from the continuum in the methoxide rather than hydroxide anion.

All the gas phase calculations have lower energy 'intermediate' structures (relative to the reactants and products) than the aqueous phase calculations. All the species had lower absolute energies in solution than in the gas phase; the effect is especially prominent in the smaller reactant and product species.

The hydroxide ion was the worst leaving group of those examined. This was followed by fluoride, with chloride being the best leaving group examined. This has been shown to be due to the effect of the substituent on the charge in the triazine ring.

The identity reactions and reactions with fluorotriazine formed Meisenheimer like σ-adduct structures, whereas the reactions with chlorotriazine (with the exception of the identity reaction) formed chloride ion-molecule complexes.

It should be noted that, practically, the slightly endothermic reactions can become favourable in solution by the utilisation of appropriate counter ions.

It has been shown that the vibrational frequencies produced by PCM and CPCM solvation approaches within Gaussian 98[46] are not sufficiently accurate in the present release, to be quoted with confidence. Analytical second derivatives for these methods, eliminating the intrinsic errors in using numerical derivatives may be available in future Gaussian packages[172] and could improve the reliability of vibrational frequencies given.

9.5) Evaluation

Density functional methods have been shown to be capable reproducing experimentally derived mechanisms in the gas phase. The results in the aqueous phase, in which most reactions occur, cannot be described reliably using the employed density functional continuum method. This reliability issue is highlighted in the inability of the continuum approaches to reproduce vibrational frequencies with confidence.

Chapter 10
Concluding remarks

The DFT studies of non-identity reactions of nucleophiles with substituted triazines in aqueous solution show two major features, of which (a) With Fluoride as a leaving group the intermediate resembles a σ-adduct with a covalent bond to both the nucleophile and leaving group (b) With Chloride as a leaving group, the intermediate resembles a chloride ion-molecule complex. This result agrees with the findings of the earlier AM1 semiempirical studies. This leads to one of two conclusions that this is either an artefact of continuum solvation models or that these reactions occur by a mechanism that is distinct from the familiar addition elimination mechanism for nucleophilic aromatic substitution of benzenoid compounds, involving a Meisenheimer complex.

The symmetrical adduct found in the DFT studies of the identity reactions of nucleophilic attack of substituted triazines in aqueous solution could, in principle, be either a intermediate or a transition structure. The standard procedure for characterisation of such a species is by determination of the vibrational frequencies; the presence of a single imaginary frequency indicates a transition structure, whereas an intermediate possesses all real frequencies.

Both CPCM and PCM within Gaussian 98 have been shown to be unreliable for the calculations of vibrational frequencies. Unfortunately even genuine intermediates may be predicted to have multiple imaginary frequencies! Although this unsatisfactory situation may be improved in future releases of the Gaussian package, at present this method should not be used for studies of reaction mechanisms in solution where there is doubt concerning the nature of any putative intermediate or transition structure.

Experimental studies have suggested that the identity reaction with pyridine as nucleophile and leaving group (pyridinolysis of 4,6-diphenoxy-1,3,5-triazin-2-yl pyridinium cation) follows a stepwise (A_N + D_N) mechanism involving an intermediate σ-adduct, whereas with 4-nitrophenolate as nucleophile and leaving group (4-nitrophenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine) proceed by means of a concerted (A_ND_N) mechanism involving only a single

transition state. The present semi-empirical calculations with continuum solvation not only show both reactions to be stepwise but also suggest the experimentally concerted reaction to have a deeper energy well for its 'intermediate'. This casts considerable doubt upon the reliability of these computational methods for studies of reaction mechanisms in solution.

The calculations for gas phase proton affinities do indicate that DFT methods in particular are capable of describing the basicity of a range of substituted pyridines and phenolates with a high accuracy. Even the semiempirical methods perform well. However, using the continuum methods to represent the effects of aqueous solvation with both DFT and semiempirical methods leads to sizeable errors in the estimation of pK_a values for the corresponding pyridinium cation and phenols. These studies were carried out with a view to establishing whether the treatment of pyridine and phenolate nucleophiles in solution was reliable. It is noteworthy that the 4-nitrophenolate anion was found to deviate significantly from the trend of the other phenolates. It is conceivable that the apparently anomalous result obtained from the calculations for the 4-nitrophenolysis of 2-(4-nitrophenoxy)-4,6-dimethoxy-1,3,5-triazine could be an artefact of this particular nucleophile.

It is possible that an improved description of these nucleophilic aromatic substitution mechanisms could be obtained by use of a method including explicit solvent molecules. For example, the effects of specific solvation upon 4-nitrophenolate and pyridine could be treated more realistically in order to reproduce the experimental observations more reliably. These methods offer much potential in the challenge of investigating reaction mechanisms in the presence of solvent.

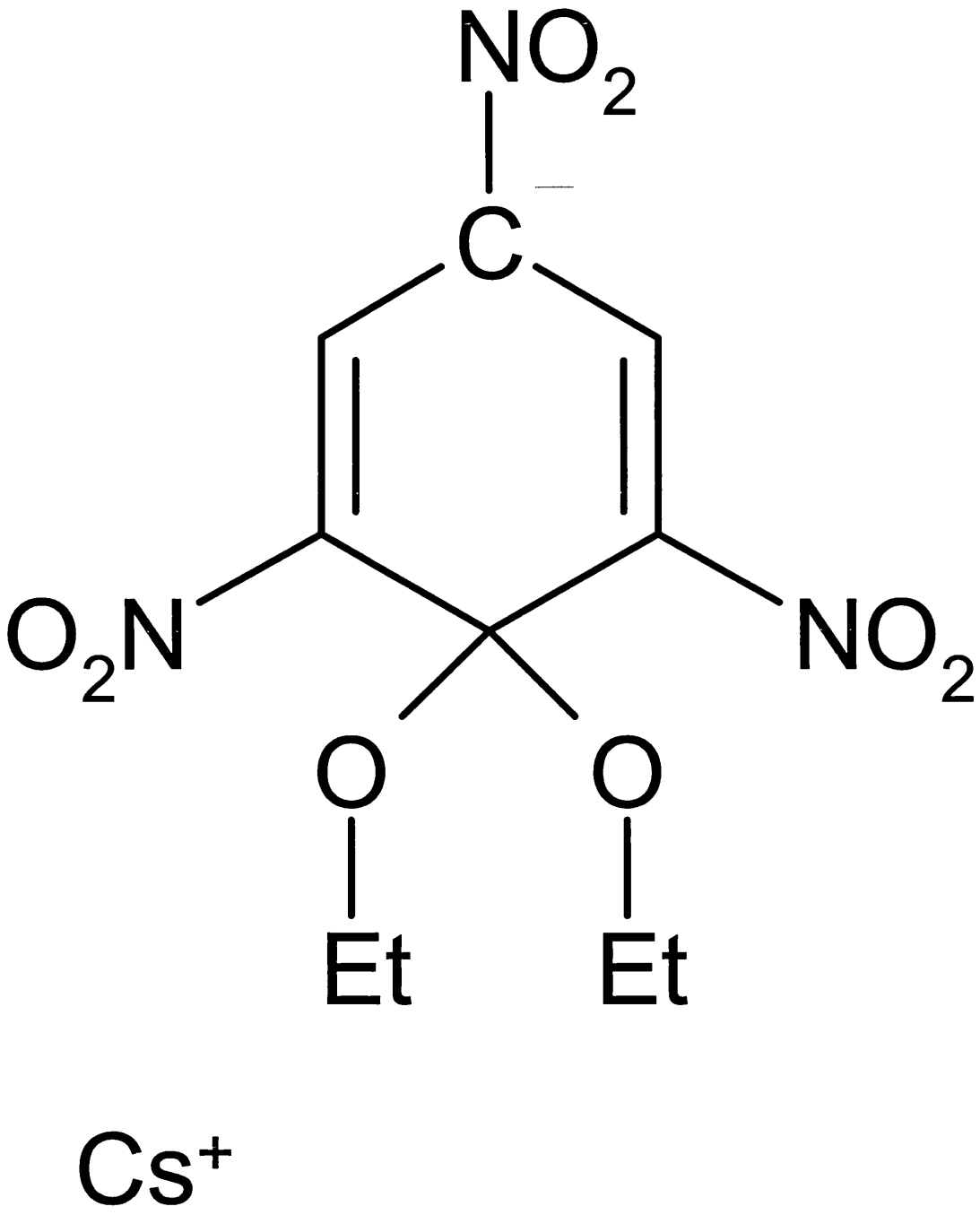
I recommend that future studies on the original problem of interest to Zeneca Specialities (now Avecia Ltd.) use other methods than the present set of continuum solvation approaches focused on here.

KMXCHY

Reference:	E.G.Kaminskaya, S.S.Gitis, A.I.Ivanova, N.V.Margolis, A.Ya.Kaminskii, N.V.Grigor'eva (1977) *Zh.Strukt.Khim.*,**18**,386
Formula:	$C_9 H_8 N_3 O_9$ 1-,K_1 1+
Compound Name:	Potassium 1-methoxycarbonyl-1-methoxy-2,4,6-trinitrocyclohexadienide

Space Group:	P21/n	***Cell:***	***a*** 18.510	***b*** 10.000	***c*** 7.640
Space Group No.:	14 ***(Å,°)***		α 90.00	β 92.00	γ 90.00

R-Factor (%):	15.0	***Temperature(K):*** 295		***Density(g/cm^3):***	1.603



TNPCSE

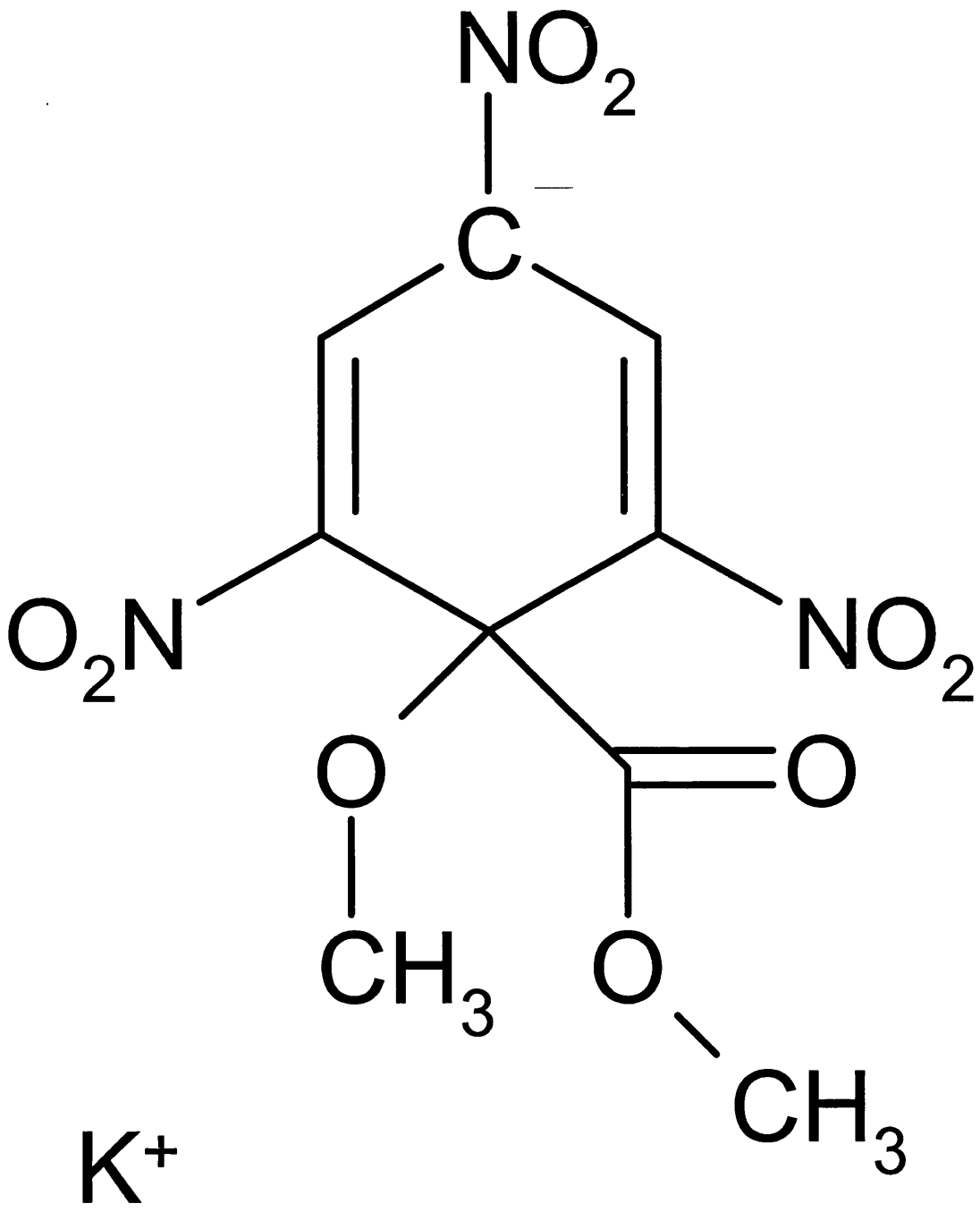
Reference: R.Destro, C.M.Gramaccioli, M.Simonetta (1968) *Acta Crystallogr.,Sect.B,***24,**1369

Formula: $C_{10}\ H_{12}\ N_3\ O_8\ 1\text{-},Cs_1\ 1+$

Compound Name: 2,4,6-Trinitrophenetole-cesium ethoxide complex

Space Group:	P21/c	***Cell:***	***a*** 15.564(3)	***b*** 10.540(20)	***c*** 19.919(1)
Space Group No.:	14 ***(Å,°)***		α 90.00	β 110.31(1)	γ 90.00

R-Factor (%): 10.2 ***Temperature(K):*** 295 ***Density(g/cm^3):*** 1.886

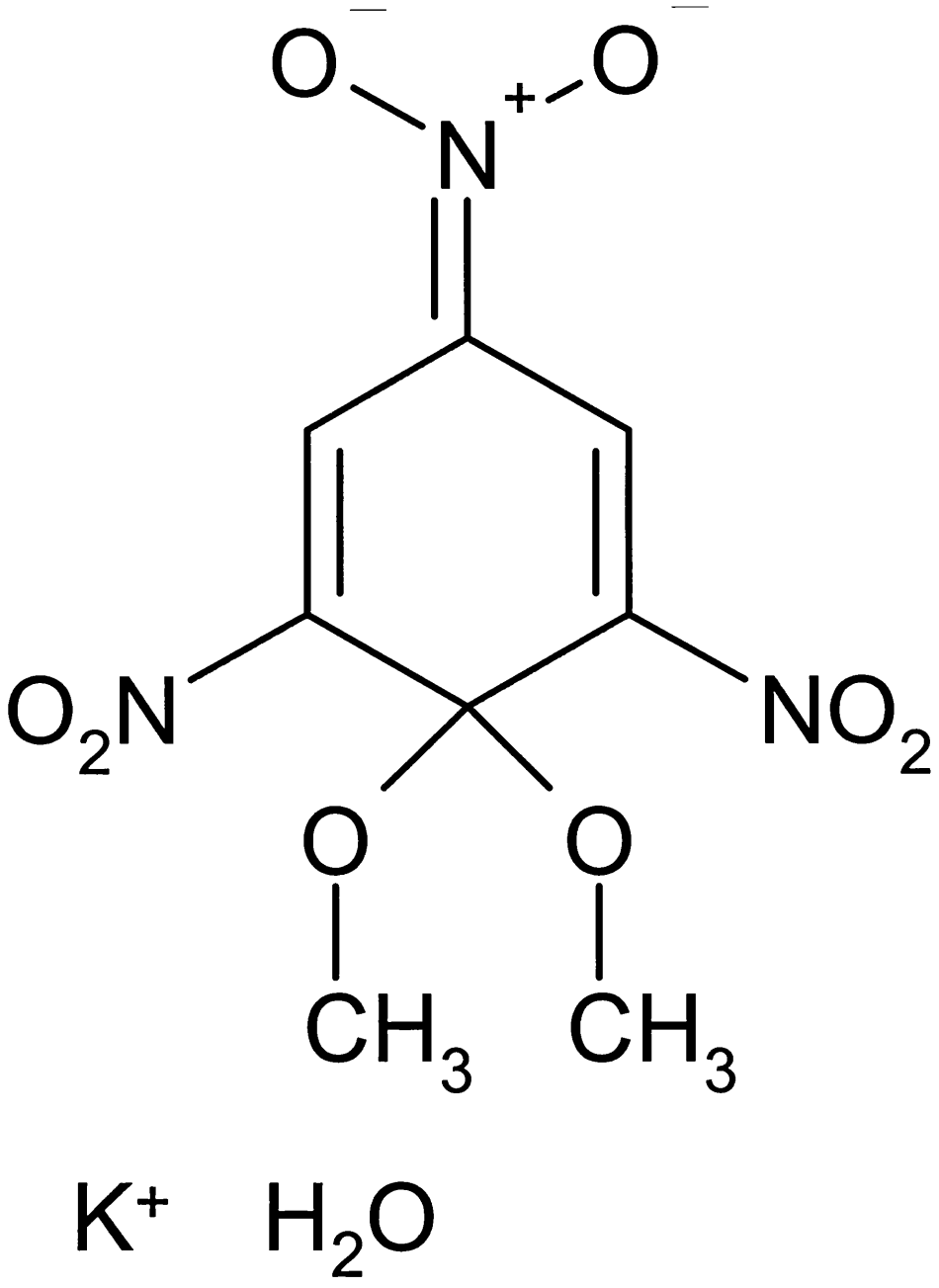


TNPTKE

Reference:	R.Destro, C.M.Gramaccioli, M.Simonetta (1968) *Acta Crystallogr.,Sect.B,***24,**1369
Formula:	$C_{10}\ H_{12}\ N_3\ O_8$ 1-,K_1 1+
Compound Name:	2,4,6-Trinitrophenetole-potassium ethoxide complex
Synonym:	Potassium Meisenheimer complex

Space Group:	P-1	***Cell:***	***a*** 14.744(1)	***b*** 10.285	***c*** 9.992
Space Group No.:	2 ***(Å,°)***		α 105.90(1)	β 104.05(1)	γ 97.15(1)

R-Factor (%):	6.4	***Temperature(K):*** 295		***Density(g/cm3):*** 1.638	



Appendix B - The total energies of the combined reactants and products along with 'intermediates' for various nucleophilic aromatic substitution reactions involving triazine, in Hartrees, as well as the differences, in kJ mol^{-1}, between the species.

Phase	Reactants		Combined Energy	Intermediate minus Reactants	Intermediate‡	Total Energy	Products minus Intermediate		Products	Combined Energy
Vacuum	tr_F	F^-	-479.404874	-155.7	tr_F_F$^-$	-479.464186#	155.7	F^-	tr_F	-479.404874
Vacuum	tr_Cl	Cl^-	-1200.137607	4.8	tr_Cl_Cl$^-$	-1200.135772#	-4.8	Cl^-	tr_Cl	-1200.137607
Vacuum	tr_OH	^-OH	-431.319527	-451.1	tr_OH_OH$^-$	-431.491363#	451.1	^-OH	tr_OH	-431.319527
Aqueous	tr_F	F^-	-479.584353	64.2	tr_F_F$^-$	-479.559889#	-64.2	F^-	tr_F	-479.584353
Aqueous	tr_Cl	Cl^-	-1200.269221	119.1	tr_Cl_Cl$^-$	-1200.223833#	-119.1	Cl^-	tr_Cl	-1200.269221
Aqueous	tr_OH	^-OH	-431.515768	64.9	tr_OH_OH$^-$	-431.491045#	-64.9	^-OH	tr_OH	-431.515768
Vacuum	tr_F	F^-	-479.404874	-155.7	tr_F_F$^-$	-479.464186#	155.7	F^-	tr_F	-479.404874
Vacuum	tr_F	^-OH	-455.336800	-236.8	tr_F_OH$^-$	-455.426998#	103.4	F^-	tr_OH	-455.387601
Vacuum	tr_F	^-OMe	-494.625490	-191.8	tr_F_OMe$^-$	-494.698552#	86.5	F^-	tr_OMe	-494.665600
Aqueous	tr_F	F^-	-479.584353	64.2	tr_F_F$^-$	-479.559889#	-64.2	F^-	tr_F	-479.584353
Aqueous	tr_F	^-OH	-455.525879	-0.6	tr_F_OH$^-$	-455.526105#	-126.4	F^-	tr_OH	-455.574242
Aqueous	tr_F	^-OMe	-494.778477	-55.2	tr_F_OMe$^-$	-494.799492#	-121.2	F^-	tr_OMe	-494.845682
Vacuum	tr_Cl	Cl^-	-1200.137607	4.8	tr_Cl_Cl$^-$	-1200.135772#	-4.8	Cl^-	tr_Cl	-1200.137607
Vacuum	tr_Cl	F^-	-839.738710	-25.3	tr_F_Cl$^-$	-839.813422*	196.1	Cl^-	tr_Cl	-839.803771
Vacuum	tr_Cl	^-OH	-815.670636	-403.9	tr_Cl_OH$^-$	-815.824518*	99.8	Cl^-	tr_OH	-815.786498
Vacuum	tr_Cl	^-OMe	-854.959326	-316.2	tr_Cl_OMe$^-$	-855.079794*	40.2	Cl^-	tr_OMe	-855.064497
Aqueous	tr_Cl	Cl^-	-1200.269221	119.1	tr_Cl_Cl$^-$	-1200.223833#	-119.1	Cl^-	tr_Cl	-1200.269221
Aqueous	tr_Cl	F^-	-839.917096	6.0	tr_F_Cl$^-$	-839.919378*	44.9	Cl^-	tr_F	-839.936478
Aqueous	tr_Cl	^-OH	-815.858622	-134.6	tr_Cl_OH$^-$	-815.909901*	-43.2	Cl^-	tr_OH	-815.926367
Aqueous	tr_Cl	^-OMe	-855.111220	-194.3	tr_Cl_OMe$^-$	-855.185242*	-33.0	Cl^-	tr_OMe	-855.197807

‡ The intermediates shown are σ-Adduct# and chloride ion-molecule complexes (Cl^-.....Tr-Y^*)

Appendix C - The gas phase vibrational frequencies for formaldehyde using the numerical and analytical HF/6-31G(d), HF/6-311+G(d,p), B3LYP/6-31G(d) and B3LYP/6-311+G(d,p) computational methods, in cm^{-1}.

	Computational	Frequencies					
HCOH	Method	B1	B2	A1	A1	A1	B2
Stepsize							
	HF/6-31G(d)	1349.4	1377.9	1671.4	1998.0	3177.9	3262.6
Analytical	HF/6-31G(d)	1351.2	1377.7	1673.8	2013.6	3172.0	3257.7
Numeric	HF/6-31G(d)	1350.6	1379.5	1674.6	2007.9	3173.3	3256.2
	HF/6-311+G(d,p)	1300.0	1358.0	1642.6	1958.5	3113.8	3201.3
Analytical	HF/6-311+G(d,p)	1353.4	1359.4	1657.1	1997.8	3152.6	3240.0
Numeric	HF/6-311+G(d,p)	1352.1	1361.5	1657.8	1994.4	3152.0	3139.0
	B3LYP/6-31G(d)	1166.7	1277.2	1556.6	1830.7	2935.5	2999.8
Analytical	B3LYP/6-31G(d)	1164.2	1265.0	1563.5	1962.2	3100.5	3163.4
Numeric	B3LYP/6-31G(d)	1163.0	1263.5	1561.6	1960.3	3096.2	3161.6
	B3LYP/6-311+G(d,p)	1146.4	1256.3	1523.5	1788.1	2903.7	2980.0
Analytical	B3LYP/6-311+G(d,p)	1172.2	1246.0	1545.9	1952.6	3080.0	3151.3
Numeric	B3LYP/6-311+G(d,p)	1170.6	1244.5	1540.6	1950.2	3077.3	3152.1

Appendix D - The calculated total energy and solvated free energy (in Hartrees), energy difference (in kJ mol^{-1}) and vibrational frequencies (in cm^{-1}) for various species.

HCN Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies π-bend	CN str	CH str
10	BLYP/6-31+G(d)	-93.407126	-93.410827	-9.7	704.6	2110.2	3376.3
1	BLYP/6-31+G(d)/CPCM(water)	-93.414636	-93.414633	0.0	726.4	2108.9	3290.0
5	BLYP/6-31+G(d)/CPCM(water)	-93.414636	-93.414634	0.0	726.4	2108.9	3290.0
10	BLYP/6-31+G(d)/CPCM(water)	-93.414636	-93.414634	0.0	726.4	2108.9	3290.0
50	BLYP/6-31+G(d)/CPCM(water)	-93.414636	-93.414633	0.0	726.4	2109.1	3290.2
100	BLYP/6-31+G(d)/CPCM(water)	-93.414636	-93.414626	0.0	726.3	2109.5	3290.7
1	BLYP/6-31+G(d)/PCM(water)	-93.414198	-93.414197	0.0	722.6	2111.4	3295.2
5	BLYP/6-31+G(d)/PCM(water)	-93.414198	-93.414197	0.0	722.5	2111.4	3295.2
10	BLYP/6-31+G(d)/PCM(water)	-93.414198	-93.414197	0.0	722.5	2111.4	3295.2
50	BLYP/6-31+G(d)/PCM(water)	-93.414198	-93.414196	0.0	722.6	2111.5	3295.4
100	BLYP/6-31+G(d)/PCM(water)	-93.414198	-93.414189	0.0	722.5	2112.0	3296.0

HCO_2^- Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6
10	BLYP/6-31+G(d)	-189.192131	-189.196221	-10.7	701.4	980.9	1274.6	1340.7	1588.0	2503.2
1	BLYP/6-31+G(d)/CPCM(water)	-189.308469	-189.308312	0.4	-1250.8	695.0	1285.8	1495.2	2775.3	3503.0
5	BLYP/6-31+G(d)/CPCM(water)	-189.308469	-189.309880	-3.7	694.9	966.1	1285.8	1495.0	1852.8	2775.2
10	BLYP/6-31+G(d)/CPCM(water)	-189.308469	-189.310294	-4.8	694.9	1092.7	1285.8	1494.9	1542.8	2775.2
50	BLYP/6-31+G(d)/CPCM(water)	-189.308469	-189.310805	-6.1	693.9	1067.6	1285.6	1345.4	1495.0	2775.4
100	BLYP/6-31+G(d)/CPCM(water)	-189.308469	-189.310864	-6.3	701.4	1036.4	1288.0	1339.3	1495.7	2776.1
1	BLYP/6-31+G(d)/PCM(water)	-189.306307	-189.309687	-8.9	-2341.0	696.3	1295.6	1533.2	1875.6	2804.4
5	BLYP/6-31+G(d)/PCM(water)	-189.306307	-189.310702	-11.5	-465.5	696.3	1295.7	1430.2	1533.1	2804.4
10	BLYP/6-31+G(d)/PCM(water)	-189.306307	-189.309406	-8.1	644.5	696.2	1295.7	1373.7	1533.1	2804.4
50	BLYP/6-31+G(d)/PCM(water)	-189.306307	-189.308739	-6.4	704.2	942.8	1296.8	1342.4	1534.9	2804.7
100	BLYP/6-31+G(d)/PCM(water)	-189.306307	-189.308674	-6.2	705.1	969.4	1298.1	1340.9	1534.7	2805.3
10	HF/6-31+G(d)	-188.208194	-188.209057	-2.3	817.8	1198.8	1494.8	1537.3	1838.9	2870.5
1	HF/6-31+G(d)/CPCM(water)	-188.331703	-188.333343	-4.3	-1686.2	802.0	1494.5	1726.6	1859.5	3122.8
5	HF/6-31+G(d)/CPCM(water)	-188.331703	-188.331817	-0.3	824.3	904.5	1524.3	1569.9	1732.5	3124.3
10	HF/6-31+G(d)/CPCM(water)	-188.331703	-188.331384	0.8	876.3	1068.1	1508.4	1540.6	1747.8	3124.1
50	HF/6-31+G(d)/CPCM(water)	-188.331703	-188.331340	1.0	817.6	1185.3	1497.6	1527.0	1734.8	3123.3
100	HF/6-31+G(d)/CPCM(water)	-188.331703	-188.331428	0.7	806.1	1177.9	1481.8	1525.3	1732.9	3123.4
1	HF/6-31+G(d)/PCM(water)	-188.329344	-188.327623	4.5	790.9	1114.0	1505.0	1764.9	2176.4	3125.9
5	HF/6-31+G(d)/PCM(water)	-188.329344	-188.328509	2.2	790.2	1313.0	1504.8	1587.0	1766.1	3125.9
10	HF/6-31+G(d)/PCM(water)	-188.329344	-188.328544	2.1	832.5	1280.5	1509.7	1545.5	1775.7	3126.5
50	HF/6-31+G(d)/PCM(water)	-188.329344	-188.328707	1.7	817.8	1241.1	1508.6	1533.7	1771.7	3126.5
100	HF/6-31+G(d)/PCM(water)	-188.329344	-188.328763	1.5	811.4	1224.7	1507.6	1533.4	1771.0	3127.0

COFH Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6
10	BLYP/6-31+G(d)	-213.755475	-213.759705	-11.1	612.5	967.6	975.5	1317.7	1812.7	3039.3
1	BLYP/6-31+G(d)/CPCM(water)	-213.761678	-213.756475	13.7	609.5	929.4	1301.1	1778.7	3025.9	4145.2
5	BLYP/6-31+G(d)/CPCM(water)	-213.761678	-213.761254	1.1	609.5	929.3	1301.2	1778.7	2047.6	3025.9
10	BLYP/6-31+G(d)/CPCM(water)	-213.761678	-213.762267	-1.5	609.3	930.0	1301.3	1667.2	1784.5	3015.2
50	BLYP/6-31+G(d)/CPCM(water)	-213.761678	-213.763354	-4.4	609.4	930.1	1146.7	1301.3	1784.6	3015.3
100	BLYP/6-31+G(d)/CPCM(water)	-213.761678	-213.763525	-4.8	609.5	929.4	1052.6	1300.9	1779.2	3026.2
1	BLYP/6-31+G(d)/PCM(water)	-213.761810	-213.756061	15.1	609.3	930.1	1301.3	1784.4	3015.3	4388.4
5	BLYP/6-31+G(d)/PCM(water)	-213.761810	-213.761168	1.7	609.3	930.0	1301.3	1784.4	2147.2	3015.2
10	BLYP/6-31+G(d)/PCM(water)	-213.761810	-213.762261	-1.2	609.3	930.0	1301.3	1667.2	1784.5	3015.2
50	BLYP/6-31+G(d)/PCM(water)	-213.761810	-213.763450	-4.3	609.4	930.1	1146.7	1301.3	1784.6	3015.3
100	BLYP/6-31+G(d)/PCM(water)	-213.761810	-213.763639	-4.8	609.5	930.2	1063.8	1301.1	1785.0	3015.6

$COHNH_3^+$ Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6
10	BLYP/6-31+G(d)	-170.155897	-170.124897	81.4	104.8	430.2	597.6	776.5	996.5	1023.4
1	BLYP/6-31+G(d)/CPCM(water)	-170.278829	-170.231187	125.1	********	505.3	507.1	753.9	1046.8	1055.3
5	BLYP/6-31+G(d)/CPCM(water)	-170.278829	-170.243495	92.8	-4435.5	505.4	510.1	754.0	1045.7	1055.1
10	BLYP/6-31+G(d)/CPCM(water)	-170.278829	-170.246064	86.0	-2759.0	505.5	518.6	754.0	1044.5	1055.2
50	BLYP/6-31+G(d)/CPCM(water)	-170.278829	-170.247437	82.4	-315.8	508.8	753.2	802.7	1047.0	1069.4
100	BLYP/6-31+G(d)/CPCM(water)	-170.278829	-170.248102	80.7	83.1	489.0	755.1	849.8	1054.6	1075.8
1	BLYP/6-31+G(d)/PCM(water)	-170.279044	-170.226767	137.2	********	-412.2	631.1	747.6	925.9	1328.0
5	BLYP/6-31+G(d)/PCM(water)	-170.279044	-170.242779	95.2	-5287.7	-209.7	642.9	731.9	952.6	1205.4
10	BLYP/6-31+G(d)/PCM(water)	-170.279044	-170.245654	87.6	-3470.5	339.5	594.5	728.6	984.1	1127.3
50	BLYP/6-31+G(d)/PCM(water)	-170.279044	-170.248151	81.1	-478.5	494.2	634.3	740.9	1038.0	1054.3
100	BLYP/6-31+G(d)/PCM(water)	-170.279044	-170.247755	82.1	183.4	498.3	740.3	814.3	1054.5	1059.3

Cl_Me_Cl⁻ Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6
10	BLYP/6-31+G(d)	-960.322179	-960.314190	21.0	-281.8	178.4	184.3	191.6	871.4	875.3
1	BLYP/6-31+G(d)/CPCM(water)	-960.406098	-960.394731	29.8	-374.2	179.1	184.8	199.0	908.4	909.9
5	BLYP/6-31+G(d)/CPCM(water)	-960.406098	-960.394732	29.8	-374.3	179.0	184.7	199.1	908.4	909.9
10	BLYP/6-31+G(d)/CPCM(water)	-960.406098	-960.394731	29.8	-374.2	179.0	184.8	199.0	908.4	909.9
50	BLYP/6-31+G(d)/CPCM(water)	-960.406098	-960.394732	29.8	-374.2	179.0	184.5	199.1	908.5	910.0
100	BLYP/6-31+G(d)/CPCM(water)	-960.406098	-960.394729	29.8	-374.0	179.0	184.7	199.1	908.6	910.0
1	BLYP/6-31+G(d)/PCM(water)	-960.405892	-960.394467	30.0	-350.9	177.8	182.0	199.6	915.0	918.1
5	BLYP/6-31+G(d)/PCM(water)	-960.405892	-960.394467	30.0	-350.9	177.8	182.0	199.6	915.0	918.1
10	BLYP/6-31+G(d)/PCM(water)	-960.405892	-960.394467	30.0	-350.9	177.8	182.0	199.6	915.0	918.1
50	BLYP/6-31+G(d)/PCM(water)	-960.405892	-960.394468	30.0	-350.9	177.8	181.8	199.6	915.0	918.1
100	BLYP/6-31+G(d)/PCM(water)	-960.405892	-960.394465	30.0	-350.8	177.8	181.9	199.6	915.2	918.3

Triazine Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6
10	BLYP/6-31+G(d)	-280.298318	-280.261675	96.2	323.5	323.5	669.3	669.3	734.5	900.7
1	BLYP/6-31+G(d)/CPCM(water)	-280.311382	-280.257256	142.1	-9728.9	321.8	442.3	661.1	664.0	828.8
5	BLYP/6-31+G(d)/CPCM(water)	-280.311382	-280.269336	110.4	-4279.7	321.8	442.3	661.3	664.1	828.8
10	BLYP/6-31+G(d)/CPCM(water)	-280.311382	-280.272142	103.0	-2962.5	321.8	442.5	661.3	664.1	828.8
50	BLYP/6-31+G(d)/CPCM(water)	-280.311382	-280.275668	93.7	-1082.0	322.1	451.3	658.9	663.4	830.1
100	BLYP/6-31+G(d)/CPCM(water)	-280.311382	-280.276215	92.3	-520.3	324.3	484.9	663.9	664.8	834.3
1	BLYP/6-31+G(d)/PCM(water)	-280.312472	-280.252948	156.3	********	314.9	521.6	664.2	678.4	830.3
5	BLYP/6-31+G(d)/PCM(water)	-280.312472	-280.268245	116.1	-5245.5	315.9	521.6	643.8	664.9	830.6
10	BLYP/6-31+G(d)/PCM(water)	-280.312472	-280.271727	107.0	-3659.5	325.6	521.9	648.3	660.9	833.5
50	BLYP/6-31+G(d)/PCM(water)	-280.312472	-280.276155	95.3	-1454.5	318.4	528.5	659.9	664.1	832.5
100	BLYP/6-31+G(d)/PCM(water)	-280.312472	-280.277029	93.0	-861.0	316.6	543.5	664.4	665.3	833.8

Triazine F F⁻ Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6=>24
10	BLYP/6-31+G(d)	-479.464186	-479.438428	67.6	47.8	274.1	363.3	467.7	467.8	473.1
1	BLYP/6-31+G(d)/CPCM(water)	-479.558435	-479.516814	109.3	-5309.2	-2369.2	-1463.0	-378.2	158.3	313.0
5	BLYP/6-31+G(d)/CPCM(water)	-479.558435	-479.526226	84.5	-2329.1	-90.0	206.5	348.2	386.8	427.1
10	BLYP/6-31+G(d)/CPCM(water)	-479.558435	-479.526954	82.6	-1601.8	235.5	245.0	355.5	416.6	439.4
50	BLYP/6-31+G(d)/CPCM(water)	-479.558435	-479.529105	77.0	-642.0	209.0	243.6	344.6	374.9	454.7
100	BLYP/6-31+G(d)/CPCM(water)	-479.558435	-479.529386	76.3	-627.5	-388.8	306.7	315.3	374.2	459.5

Triazine Cl Cl⁻ Stepsize	Computational Method	Total Energy	Free Energy	ΔE	Frequencies 1	2	3	4	5	6=>24
10	BLYP/6-31+G(d)	-1200.135772	-1200.114868	54.9	-149.7	123.2	193.0	193.6	258.7	298.2
1	BLYP/6-31+G(d)/CPCM(water)	-1200.224386	-1200.172486	136.2	-3392.7	-2689.4	-1433.9	-1228.9	-379.7	189.0
5	BLYP/6-31+G(d)/CPCM(water)	-1200.224386	-1200.189213	92.3	-1480.3	-1013.9	-497.6	-161.9	191.2	228.6
10	BLYP/6-31+G(d)/CPCM(water)	-1200.224386	-1200.194071	79.6	-682.3	-599.1	-289.9	102.5	182.4	224.7
50	BLYP/6-31+G(d)/CPCM(water)	-1200.224386	-1200.187236	97.5	-7221.8	-4721.7	-2889.1	-176.6	210.2	276.7
100	BLYP/6-31+G(d)/CPCM(water)	-1200.224386	-1200.198972	66.7	-249.7	-239.9	46.5	185.7	226.6	295.6

Chapter 11

References

1. Buttar, D., *Molecular modelling of reactive dye units, Internal technical report*, . 1997, Zeneca Specialities.
2. Buttar, D.D., R., *Molecular modelling study of the hydrolysis of reactive dyes, Internal technical report*, . 1996, Zeneca Specialities.
3. Hoskin, P., *9 month report*, . 1997, University of Bath.
4. Lavery, A. and J. Provost, *Color-media interactions in ink jet printing, A meeting abstract*, . 1998, Zeneca Specialities.
5. Rys, P. and H. Zollinger, *Chapter 6 - Reactive dye-fibre systems*, in *The theory of coloration of textiles (2nd edition)*, A. Johnson, Editor. 1989, The Society of dyers and colourists. p. 428.
6. Dawson, T.L., J. Soc. Dyers Colourists, 1963. **80**: p. 4130.
7. Meisenheimer, J., Justus Liebigs Ann. Chem., 1902. **323**: p. 205.
8. Jackson, C.J. and F.H. Gazzolo, J. Am. Chem. Soc., 1900. **23**: p. 376.
9. Terrier, F., *Rate and Equilibrium Studies in Jackson-Meisenheimer Complexes.* Chem. Rev., 1982. **82**(2): p. 78.
10. Strauss, M.J., *Carboanion additions and cylizations involving anioinic sigma complexes. Meta bridging reactions of aromatics.* Acc. Chem. Res., 1974. **7**: p. 181.
11. Strauss, M.J., *Anionic sigma complexes.* Chem. Rev., 1970. **70**(6): p. 667.
12. Crampton, M.R., *Meisenhiemer complexes.* Adv. Phys. Org. Chem., 1969. **7**: p. 211.
13. Illuminati, G. and F. Stegel, *The Formation of Anionic Sigma-Adducts From Heteroaromatic-Compounds - Structures, Rates, and Equilibria.* Adv. Heterocyclic. Chem., 1983. **34**: p. 305.
14. Artamkina, G.A., M.P. Egorov, and I.P. Beletskaya, *Some Aspects of Anionic Sigma-Complexes.* Chem. Rev., 1982. **82**(4): p. 427.
15. Crampton, M.R. and V. Gold, *Reactions of aromatic nitro-compounds in alkaline media. Part IX. Nuclear Magnetic Resonance spectra of Meisenheimer complexes.* J. Chem. Soc., 1964: p. 4293.
16. Servis, K.L., *Nuclear Magnetic Resonance studies of Meisenheimer complexes.* J. Am. Chem. Soc., 1967. **89**(6): p. 1508.
17. Foster, R. and C.A. Fyfe, *Interaction of electron acceptors with bases - XVI - Proton magnetic resonance spectra of Meisenheimer compounds.* Tetrahedron, 1965. **21**: p. 3363.
18. Fyfe, C.A., S.W.H. Damji, and A. Koll, *Low-temperature flow NMR investigation of the transient intermediate in the nucleophilic aromatic substitution of 2,4,6-trinitroanisole by n-butylamine.* J. Am. Chem. Soc., 1979. **101**(4): p. 951.
19. Olah, G.A. and H. Mayr, *Carboanions. 2. Corbon-13 Nuclear Magnetic Resonance study of Meisenheimer complexes and their charge distribution pattern.* J. Org. Chem., 1976. **41**(21): p. 3448.

20. Destro, R., C.M. Gramaccioli, and M. Simonetta, *The crystal and molecular structure of the complexes of 2,4,6-trinitrophenotole with caesium or potassium ethoxide (Meisenheimer salts).* Acta Crystallogr., 1968. **B24**: p. 1369.
21. Messmer, G.G. and G.J. Palenik, *The crystal structure of a Meisenheimer complex: the potassium methoxide adduct of 4-methoxy-5,7-dinitrobenzofurazan.* Chem. Commun., 1969: p. 470.
22. Hunter, A., *et al.*, *A Single Transition-State in Nucleophilic Aromatic-Substitution - Reaction of Phenolate Ions With 2-(4-Nitrophenoxy)-4,6-Dimethoxy- 1,3,5-Triazine in Aqueous-Solution.* J. Chem. Soc., Perkin Trans. 2, 1993(10): p. 1703.
23. Pople, J.A. and D.L. Beveridge, *Approximate Molecular Orbital Theory.* 1970, New York: MacGraw-Hill.
24. Heilbronner, E. and H. Bock, *The HMO Model and its Application.* 1976: Wiley-Interscience.
25. Szabo, A. and N.S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced-Electronic Structure Theory.* 1982, New York: MacMillen.
26. McQuarrie, D.A., *Quantum Chemistry.* 1983: Oxford University Press.
27. Grant, G.H. and W.G. Richards, *Computational Chemistry.* 1995: Oxford University Press.
28. Atkins, P.W. and R.S. Friedman, *Molecular Quantum Mechanics.* Third edition ed. 1997: Oxford University Press.
29. Leach, A.R., *Molecular Modelling - Principles and applications.* Longman. 1996.
30. Møller, C. and M.S. Plesset, Phys. Rev., 1934. **46**: p. 618.
31. Almlöf, J., K. Faegri, and K. Korsell, J. Comput. Chem., 1982. **3**: p. 385.
32. Slater, J.C., Phys. Rev., 1930. **36**: p. 57.
33. Huzinaga, Z., J. Chem. Phys., 1965. **42**: p. 1293.
34. Hehre, W.J., R.F. Stewart, and J.A. Pople, J. Chem. Phys., 1969. **51**: p. 2657.
35. Ziegler, T., *Approximate Density Functional Theory As a Practical Tool in Molecular Energetics and Dynamics.* Chem. Rev., 1991. **91**(5): p. 651.
36. Andzelm, J. and E. Wimmer, *Density functional Gaussian-type-orbital approach to molecular geometries, vibrations, and reaction energies.* J. Chem. Phys., 1992. **96**(2): p. 1280.
37. Hobza, P., J. Sponer, and M. Polasek, *H-Bonded and Stacked Dna-Base Pairs - Cytosine Dimer - an Ab-Initio 2nd-Order Moller-Plesset Study.* J. Am. Chem. Soc., 1995. **117**(2): p. 792.
38. Dewar, M.J.S. and W. Thiel, J. Am. Chem. Soc., 1977. **99**: p. 4907.
39. Dewar, M.J.S. and W. Thiel, J. Am. Chem. Soc., 1977. **99**: p. 4899.
40. Dewar, M.J.S., *et al.*, *The Development and Use of Quantum-Mechanical Molecular-Models .76. Am1 - a New General-Purpose Quantum-Mechanical Molecular-Model.* J. Am. Chem. Soc., 1985. **107**(13): p. 3902.
41. Bingham, R.C., M.J.S. Dewar, and D.W. Lo, J. Am. Chem. Soc., 1975. **97**: p. 1294.
42. Bingham, R.C., M.J.S. Dewar, and D.W. Lo, J. Am. Chem. Soc., 1975. **97**: p. 1285.
43. Bingham, R.C., M.J.S. Dewar, and D.W. Lo, J. Am. Chem. Soc., 1975. **97**: p. 1302.
44. Bingham, R.C., M.J.S. Dewar, and D.W. Lo, J. Am. Chem. Soc., 1975. **97**: p. 1307.
45. Stewart, J.P.P., J. Comput. Chem., 1989. **10**: p. 209.

46. Frisch, M.J., *et al.*, *Gaussian 98*, . 1998, Gaussian, Inc.: Pittsburgh, PA.
47. Cramer, C.J. and D.G. Truhlar, *Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics.* Chem. Rev., 1999. **99**: p. 2161.
48. Tapia, O. and J. Bertrán, eds. *Solvent Effects and Chemical Reactivity.* Understanding Chemical Reactivity. 1996, Kluwer Academic Publishers: Dordrecht, Boston, London.
49. Tomasi, J. and M. Persico, *Molecular-Interactions in Solution - an Overview of Methods Based On Continuous Distributions of the Solvent.* Chem. Rev., 1994. **94**(7): p. 2027.
50. Born, M., *Volumen und Hydratationswärme der Ionen.* Zeitshift für Physik, 1920. **1**: p. 45.
51. Onsager, L., *Electric Moments of Molecules in Liquids.* J. Am. Chem. Soc., 1936. **58**: p. 1486.
52. Klamt, A. and G. Schuurmann, *COSMO - a New Approach to Dielectric Screening in Solvents With Explicit Expressions For the Screening Energy and Its Gradient.* J. Chem. Soc., Perkin Trans. 2, 1993(5): p. 799.
53. Hawkins, G.D., *et al.*, *AMSOL version 6.1*, . 1997, Oxford Molecular.
54. Liotard, D.A., *et al.*, *Improved Methods For Semiempirical Solvation Models.* J. Comput. Chem., 1995. **16**(4): p. 422.
55. Giesen, D.J., C.J. Cramer, and D.G. Truhlar, J. Phys. Chem., 1995. **99**: p. 7137.
56. Giesen, D.J., *et al.*, *General Semiempirical Quantum-Mechanical Solvation Model For Nonpolar Solvation Free-Energies - N-Hexadecane.* J. Am. Chem. Soc., 1995. **117**(3): p. 1057.
57. Miertuš, S., E. Scrocco, and J. Tomasi, J. Chem. Phys., 1981. **55**: p. 117.
58. Cammi, R. and J. Tomasi, J. Chem. Phys., 1994. **100**: p. 7495.
59. Barone, V. and M. Cossi, *Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model.* J. Phys. Chem. A, 1998. **102**(11): p. 1995.
60. Hoski, H., *et al.*, J. Chem. Phys., 1987. **87**: p. 1107.
61. Cossi, M., B. Mennucci, and J. Tomasi, Chem. Phys. Lett., 1994. **165**: p. 228.
62. Bonaccorsi, R., *et al.*, J. Theo. Chem. Acta, 1990. **162**: p. 213.
63. Bonaccorsi, R., *et al.*, J. Chem. Phys., 1990. **143**: p. 245.
64. Cramer, C.J. and D.G. Truhlar, *General Parameterized Scf Model For Free-Energies of Solvation in Aqueous-Solution.* J. Am. Chem. Soc., 1991. **113**(22): p. 8305.
65. Cramer, C.J. and D.G. Truhlar, Science, 1992. **256**: p. 213.
66. Cramer, C.J. and D.G. Truhlar, *Pm3-Sm3 - a General Parameterization For Including Aqueous Solvation Effects in the Pm3 Molecular-Orbital Model.* J. Comput. Chem., 1992. **13**(9): p. 1089.
67. Storer, J.W., *et al.*, *Structure and Reactivity in Aqueous Solution*, , C.J. Cramer and D.G. Truhlar, Editors. 1994, American Chemical Society: Washington DC. p. 24.
68. Giesen, D.J., *et al.*, *A universal organic solvation model.* J. Org. Chem., 1996. **61**(25): p. 8720.
69. Maskill, H., *The physical basis of organic chemistry.* 1985, Oxford: Oxford science publications. 490.
70. Gandour, R.D. and R.L. Schowen, *Transition states of biochemical processes.* 1978, New York and London: Plenum Press.
71. Barnes, J.A., J. Wilkie, and I.H. Williams, J. Chem. Soc., Faraday Trans., 1994. **90**(1709).

72. McIver, J.W., Jr. and A. Komornicki, J. Am. Chem. Soc., 1972. **94**: p. 2625.
73. Cooper, D.L., M.A. Robb, and I.H. Williams, Chem. Brit., 1994: p. 1085.
74. Stewart, J.J.P., *Mopac 93.* J. Computer-Aided Mol. Des., 1990. **4**: p. 1.
75. Dewar, M.J.S., E.F. Healy, and J.J.P. Stewart, *Location of Transition States in Reaction Mechanisms.* J. Chem. Soc., Faraday Trans. 2, 1984. **80**: p. 227.
76. Young, D.C., *Computational Chemistry - A Practical Guide for Applying Techniques to Real World Problems.* 2001: Wiley-Interscience. 381.
77. Chamberlin, R.A., M.R. Crampton, and I.A. Robotham, *Kinetics of the Reactions of Alkyl 2,4,6-Trinitrophenyl Ethers With Amines in Dimethyl-Sulfoxide - the Mechanism of Base Catalysis.* J. Chem. Res., Synop., 1994(11): p. 408.
78. Bacaloglu, R., *et al.*, *Mechanism of reaction of Hydroxide Ion with Dinitrochlorobenzenes.* J. Am. Chem. Soc., 1991. **113**(1): p. 238.
79. Buncel, E., *et al.*, *Regioselectivity in the Reaction of Ambident Phenoxide Ion and Methoxide and Hydroxide Ions With 2,4,6-Trinitroanisole - Kinetic and Thermodynamic Control.* J. Am. Chem. Soc., 1992. **114**(14): p. 5610.
80. Bacaloglu, R., C.A. Bunton, and F. Ortega, *Single-Electron Transfer in Aromatic Nucleophilic Substitution Reactions of 1 Substituted-2,4-Dinitronaphthalenes with Hydroxide Ion.* J. Am. Chem. Soc., 1988. **110**(11): p. 3512.
81. Bacaloglu, R., C.A. Bunton, and F. Ortega, *Multistep Reaction Analysis - A Numerical Approach Based on Relaxation Theory.* Int. J. Chem. Kinet., 1988. **20**(3): p. 195.
82. Bacaloglu, R., C.A. Bunton, and F. Ortega, *Single-Electron Transfer in Aromatic Nucleophilic Addition and Substitution in Aqueous-Media.* J. Am. Chem. Soc., 1988. **110**(11): p. 3503.
83. Bacaloglu, R., *et al.*, *Proton-Exchange and Nuclear Magnetic-Resonance Line Broadening in Arotmatic Nucleophilic-Addition and Substitution.* J. Am. Chem. Soc., 1988. **110**(11): p. 3503.
84. Bacaloglu, R., C.A. Bunton, and G. Cerichelli, *Intermediates in Nucleophilic Aromatic Substitution.* J. Am. Chem. Soc., 1987. **109**(2): p. 621.
85. Dotterer, S.K. and R.L. Harris, *MNDO Study of Nucleophilic Aromatic Substitution.* J. Org. Chem., 1988. **53**(4): p. 777.
86. Bacaloglu, R., C.A. Bunton, and F. Ortega, *Interaction of Nitroarenes With Hydroxide Ion - an Am1 Molecular- Orbital Treatment.* J. Am. Chem. Soc., 1989. **111**(3): p. 1041.
87. Bacaloglu, R., *et al.*, *Single-Electron Transfer in Aromatic Nucleophilic Substitution.* J. Am. Chem. Soc., 1992. **114**(20): p. 7708.
88. Simkin, B.Y., V.I. Minkin, and M.N. Glukhovtsev, *Theoretical Study of Mechanisms of Aromatic Nucleophilic Substitution in the Gas Phase.* THEOCHEM, 1993. **103**(1-2): p. 123.
89. Buncel, E., R.M. Tarkka, and J.M. Dust, *Am1 Studies On the Stabilities of Anionic Sigma-Complex Regioisomers - Thermodynamics of Regioselectivity in the Reaction of Methide, Methoxide, and Hydroxide Anions With Electron-Deficient Aromatics.* Can. J. Chem., 1994. **72**(7): p. 1709.
90. Renfrew, A.H.M., *et al.*, *Nucleophilic Aromatic-Substitution in Heterocycles - Alcoholysis and Hydrolysis of 2-Anilino-4,6-Dichloro-1,3,5-Triazines.* J. Chem. Soc., Perkin Trans. 2, 1994(12): p. 2389.

91. Renfrew, A.H.M., *et al.*, *Timing of Bonding Changes in Fundamental Reactions in Solutions - Pyridinolysis of a Triazinylpyridinium Salt.* J. Chem. Soc., Perkin Trans. 2, 1994(12): p. 2383.
92. Shakes, J., *et al.*, *Concerted displacement mechanisms at trigonal carbon: The aminolysis of 4-aryloxy-2,6-dimethoxy-1,3,5-triazines.* J. Chem. Soc., Perkin Trans. 2, 1996. **8**: p. 1553.
93. Bentley, T.W., *et al.*, *Homogeous Models for the Chemical Selectivity of Reactive Dyes on Cotton - Developement of Procedures and Choice of Model.* J. Soc. Dyers Colourists, 1995. **111**(9): p. 288.
94. Ramesh, A., B.S. Sundar, and P.S.R. Murti, *Aromatic Nucleophilic Substitution Reactions - Kenetics of Substitution of 2-Chloropyrimimidine and 2,4,6-trichloro-1,3,5-triazine by various nucleophiles.* J. Indian Chem. Soc., 1995. **72**(10): p. 697.
95. Hawkins, G.D., *et al.*, *AMSOL version 5.4.1.* QCPE Bulletin 16, 1996: p. 11.
96. Dieter, K.M. and J.J.P. Stewart, *Calculation of Vibrational Frequencies using Molecular Trajectories.* THEOCHEM, 1988. **40**: p. 143.
97. Yamabe, S., T. Minato, and Y. Kawabata, *The Importance of the Sigma-Star-Pi-Star Orbital Mixing for the Nucleophilic Displacement of the Unsaturated Carbon.* Can J. Chem., 1984. **62**(2): p. 235.
98. Guest, M.F., *et al.*, *Computational Chemistry on the FPS-X64 Scientific Computers - Experience of Single-Processor and Multiprocessor Systems.* Theor. Chim. Acta, 1987. **71**(2-3): p. 117.
99. Lias, S.G., *et al.*, *Gas-Phase Ion and Neutral Thermochemistry.* J. Phys. Chem. Ref. Data Suppl., 1988. **17**(1): p. 1.
100. *Handbook of Chemistry and Physics*, ed. W. R.C. and M.J. Astle. 1984: CRC Press.
101. Cullum, N.R., *et al.*, *Effective Charge on the Nucleophile and Leaving Group During the Stepwise Transfer of the Triazinyl Group Between Pyridines in Aqueous-Solution.* J. Am. Chem. Soc., 1995. **117**(36): p. 9200.
102. Hunter, A., *et al.*, J. Chem. Soc., Perkin Trans. 2, 1994: p. 2389.
103. Cullum, N.R., *et al.*, *The aminolysis and hydrolysis of N-(4,6-diphenoxy-1,3,5-triazin-2-yl) substituted pyridinium salts: Concerted displacement mechanism.* J. Chem. Soc., Perkin Trans. 2, 1996(8): p. 1559.
104. Hunter, A., *et al.*, J. Chem. Soc., Perkin Trans. 2, 1994: p. 2383.
105. Hunter, A., *et al.*, J. Am. Chem. Soc., 1995. **117**: p. 5484.
106. Robotham, I.A., *Mechanistic Studies of aromatic substitutions with aniline and phenoxide nucleophiles*, in *Department of Chemistry*. 1997, University of Durham.
107. Chambers, C.C., *et al.*, *Model for aqueous solvation based on class IV atomic charges and first solvation shell effects.* J. Phys. Chem., 1996. **100**(40): p. 16385.
108. **Schaftenaar, G. and J.H. Noordik, *Molden: A Pre- and Post-Processing Program for Molecular and Electronic Structures*. J. Computer-Aided Mol. Des.**, 2000. **14**(2): p. 123.
109. Becke, A.D., *Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior.* Phys. Rev. A, 1988. **38**(6): p. 3098.

110. Lee, C.T., W.T. Yang, and R.G. Parr, *Development of the Colle-Solvetti Correlation-Energy Formula into a Functional of the Electron Density.* Phys. Rev. B, 1988. **37**(2): p. 785.

111. McMahon, T.B. and P. Kebarle, *Intrinsic acidities of substituted phenols and benzoic acids determined by gas-phase proton-transfer equilibria.* J. Am. Chem. Soc., 1977. **99**(7): p. 2222-2230.

112. Aue, D.H., *et al.*, *Relationships Between the Thermodynamics of Protonation in the Gas and Aqueous Phase For 2-Substituted, 3-Substituted, and 4-Substituted Pyridines.* J. Am. Chem. Soc., 1991. **113**(5): p. 1770.

113. Voets, R., *et al.*, *Theoretical-Study of the Proton Affinities of 2-Monosubstituted, 3-Monosubstituted, and 4-Monosubstituted Pyridines in the Gas-Phase By Means of Mindo/3, Mndo, and Am1.* J. Comput. Chem., 1989. **10**(4): p. 449.

114. Szafran, M. and J. Koput, *Pm3 Study of the Proton Affinities of 2-Monosubstituted, 3-Monosubstituted, and 4-Monosubstituted Pyridines in the Gas-Phase.* J. Comput. Chem., 1991. **12**(6): p. 675.

115. Voets, R., *et al.*, *Theoretical-Study of the Proton Affinities of 2-Monosubstituted, 3-Monosubstituted, and 4-Monosubstituted Phenolate Ions in the Gas- Phase By Means of Mindo/3, Mndo, and Am1.* J. Comput. Chem., 1990. **11**(3): p. 269.

116. Caronna, T. and B.M. Vittimberga, *Correlation of 2-Substituted, 3-Substituted, 4-Substituted and Disubstituted Pyridine Gas-Phase Proton Affinities With Abinitio Calculated Energies At the Sto-3g Basis Set Level.* J. Heterocycl. Chem., 1992. **29**(4): p. 787.

117. Smith, B.J. and L. Radom, *Calculation of Proton Affinities Using the G2(Mp2,Svp) Procedure.* J. Phys. Chem., 1995. **99**(17): p. 6468.

118. *JANEF Thermodynamic Tables.* 2nd ed, ed. D.R. Stull and H. Prophet. 1971, Washington D.C.

119. Atkins, P. and J. De Paula, *Physical Chemistry.* Seventh ed. 2001, Oxford: Oxford University Press.

120. Helgaker, T., Chem. Phys. Lett., 1991. **782**: p. 503.

121. Culot, P., *et al.*, Theor. Chim. Acta, 1992. **82**: p. 189.

122. Banerjee, A., *et al.*, J. Phys. Chem., 1985. **89**: p. 52.

123. Baker, J., J. Comput. Chem., 1986. **7**: p. 385.

124. Urban, J.J., R.L. Vontersch, and G.R. Famini, *Effect of Fluorine Substitution On Phenol Acidities in the Gas-Phase and in Aqueous-Solution - a Computational Study Using Continuum Solvation Models.* J. Org. Chem., 1994. **59**(18): p. 5239.

125. Szafran, M., *et al.*, *Reconsideration of Solvent Effects Calculated By Semiempirical Quantum Chemical Methods.* J. Comput. Chem., 1993. **14**(3): p. 371.

126. Kebarle, P. and T.B. McMahon, *Intrinsic Acidities of Substituted Phenols and Benzoic Acids Determined by Gas Phase Proton Transfer Equilibria.* J. Am. Chem. Soc., 1977. **99**: p. 2222.

127. Fujio, M. and R.T. McIver, Jr., *Effects on the acidities of phenols from specific substituent-solvent interactions. Inherent substituent parameters from gas phase acidities.* J. Am. Chem. Soc., 1981. **103**: p. 4017.

128. Hernandez-Gill, N., W.E. Wentworth, and E.C.M. Chen, *Electron affinities of fluorinated phenoxy radicals.* J. Phys. Chem., 1984. **88**: p. 6181.

129. Gunion, R.F., *et al.*, *Ultraviolet Photoelectron Spectroscopy of the Phenide, Benzyl, and Phenoxide Anions.* Int. J. Mass Spectrom. Ion Proc., 1992. **117**: p. 601.
130. Hunter, E.P.L. and S.G. Lias, *Evaluated gas phase basicities and proton affinities of molecules: An update.* J. Phys. Chem. Ref. Data, 1998. **27**(3): p. 413.
131. Scott, A.P. and L. Radom, *Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Moller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempricial Scale Factors.* J. Phys. Chem., 1996. **100**: p. 16502.
132. Cramer, C.J. and D.G. Truhlar, eds. *Structure and reactivity in aqueous solution.* . 1994, ACS symposium series 568: Washington D.C.
133. Truong, T.N. and E.V. Stefanovich, *Hydration Effects On Reaction Profiles - an Ab-Initio Dielectric Continuum Study of the S(N)2 Cl-+Ch3cl Reaction.* J. Phys. Chem., 1995. **99**(40): p. 14700.
134. Klamt, A., *et al.*, *Refinement and parametrization of COSMO-RS.* J. Phys. Chem. A, 1998. **102**(26): p. 5074.
135. Klamt, A., *Conductor-Like Screening Model For Real Solvents - a New Approach to the Quantitative Calculation of Solvation Phenomena.* J. Phys. Chem., 1995. **99**(7): p. 2224.
136. Foresman, J.B., *et al.*, J. Phys. Chem., 1996. **100**: p. 16098.
137. Dillet, V., D. Rinaldi, and J.L. Rivail, J. Phys. Chem., 1996. **104**: p. 9437.
138. Dillet, V., D. Rinaldi, and J.L. Rivail, J. Phys. Chem., 1994. **94**: p. 5034.
139. Moliner, V., A.J. Turner, and I.H. Williams, *Transition-state structural refinement with GRACE and CHARMM: Realistic modelling of lactate dehydrogenase using a combined quantum/classical method.* Chem. Commun., 1997(14): p. 1271.
140. Still, W.C., *et al.*, J. Am. Chem. Soc., 1990. **112**: p. 6127.
141. Schüürmann, G., *Modelling pK(a) of carboxylic acids and chlorinated phenols.* QSAR, 1996. **15**(2): p. 121.
142. Schüürmann, G., *Quantum chemical analysis of the energy of proton transfer from phenol and chlorophenols to H2O in the gas phase and in aqueous solution.* J. Chem. Phys., 1998. **109**(21): p. 9523.
143. Schüürmann, G., *et al.*, *Prediction of the pK(a) of carboxylic acids using the ab initio continuum-solvation model PCM-UAHF.* J. Phys. Chem. A, 1998. **102**(33): p. 6706.
144. Chen, J.I. and A.D.J. MacKerell, *Computation of the influence of chemical substitution on the pK_a of pyridine using semiempirical and ab initio methods.* Theor. Chem. Acc., 2000. **103**: p. 483.
145. Brondi, A., J. Chem. Phys., 1964. **68**: p. 441.
146. Liptak, M.D. and G.C. Shields, *Accurate pK_a Calculations for Carboxlic Acids Using Complete Basis Set and Gaussian-n ModelsCombined with CPCM Continuum Solvation Methods.* J. Am. Chem. Soc., 2001. **123**: p. 7314.
147. Tissandier, M.D., *et al.*, *The proton's absolute aqueous enthalpy and Gibbs free energy of solvation from cluster-ion solvation data.* J. Phys. Chem. A, 1998. **102**(40): p. 7787.
148. Ben-Naim, A., J. Phys. Chem., 1978. **82**: p. 792.
149. Liotta, C.L., E.M. Perdue, and J.H.P. Hopkins, *Thermodynamics of acid-base equilibria. VI. Ionization of substituted pyrinium ions.* J. Am. Chem. Soc., 1974. **96**(23): p. 7308.

150. Parsons, G.H., C.H. Rochestr, and C.E.C. Wood, *Effect of 4-substitution on the thermodynamics of hydration of phenol and the phenoxide anion.* J. Chem. Soc. B, 1971: p. 533.
151. Kortum, G., W. Vogel, and K. Andrussow, *Dissociation constants of organic acids in aqueous solution.* 1961, London: Butterworths.
152. Fernandez, L.P. and L.G. Hepler, J. Phys. Chem., 1959(81): p. 1783.
153. Hepler, L.G. and W.F. O'Hara, J. Phys. Chem., 1961. **65**: p. 811.
154. O'Hara, W.F. and L.G. Hepler, J. Phys. Chem., 1961. **65**: p. 2107.
155. Millera, F.J., J.C. Aliluwalia, and L.G. Hepler, J. Chem. Eng. Data, 1964. **9**(192): p. 319.
156. Aue, D.H., H.M. Webb, and M.T. Bowers, J. Am. Chem. Soc., 1976. **98**: p. 311.
157. Weselucha-Birczynska, A. and M. Ciechanowicz-Rutkowska, *Experimental and calculated structure of vibrational spectra of cinchonine.* THEOCHEM, 2000(555): p. 391.
158. Majer, V. and V. Svoboda, *Enthalpies of Vaporization of Organic Compounds: A Critical Review and Data Compilation*, . 1985, Blackwell Scientific Publications: Oxford. p. 300.
159. Fischer, A., W.J. Galloway, and J. Vaughan, *Structure and reactivity in the pryidine series. Part I. Acid dissociation constants of pyridinium ions.* J. Chem. Soc., 1964: p. 3591.
160. Elderfield, R.C., *Polycyclic compounds containing two hetero atoms in different rings. Five- and six-membered heterocycles containing three hetero atoms and their benzo derivatives.* Heterocyclic compounds. Vol. 7. 1961, New York: Wiley. 878.
161. Boulton, A.J. and A. Mc Killop, eds. *Six-membered rings with oxygen, sulphur or two or more nitrogen atoms.* Comprehensive heterocyclic chemistry the structure, reactions, synthesis, and uses of heterocyclic compounds, ed. A.R. Katritzky and C.W. Rees. Vol. 3, 2B. 1984, Pergamon: Oxford.
162. Albert, A., *Heterocyclic Chemistry an Introduction.* 2nd ed. 1968, London: The Athlone Press. 547.
163. Allen, F.H. and O. Kennard, *3D Search and Research Using the Cambridge Structural Database.* Chem. Design Automation News, 1993. **8**(1): p. 1,31-37.
164. Drew, M.G.B., *et al.*, *2,4,6-Tris(2-pyridyl)-1,3,5-triazine.* Acta Crystallogr., Sect. C: Cryst. Struct. Commun., 1998. **54**: p. 985.
165. Kupka, T., I.P. Gerothanassis, and I.N. Demetropoulos, *Density functional study of a model amide. Prediction of formamide geometry, dipole moment, IR harmonic vibration nu(C=O) and GIAO NMR shieldings.* THEOCHEM, 2000. **531**: p. 143.
166. Mulliken, R.S., *Electron population analysis on LCAO-MO molecular wave functions. I.* J. Chem. Phys., 1955. **23**(10): p. 1833.
167. Mulliken, R.S., *Electron population analysis on LCAO-MO molecular wave functions. II. Overlap populations, bond orders, and covalent bond energies.* J. Chem. Phys., 1955. **23**(10): p. 1841.
168. Mulliken, R.S., *Electron population analysis on LCAO-MO molecular wave functions. III. Effects of hybridization on overlap and gross AO populations.* J. Chem. Phys., 1955. **23**(12): p. 2338.
169. Mulliken, R.S., *Electron population analysis on LCAO-MO molecular wave functions. IV. Bonding and antibonding in LCAO and valence-bond theories.* J. Chem. Phys., 1955. **23**(12): p. 2343.

170. Bartmess, J.E., J.A. Scott, and J.R.T. McIver, *Scale of acidities in the gas-phase from methanol to phenol.* J. Am. Chem. Soc., 1979. **101**(20): p. 6046.

171. Shimanouchi, T., *Tables of Molecular Vibrational Frequencies Consolidated Volume I.* 1972: National Bureau of Standards. 1.

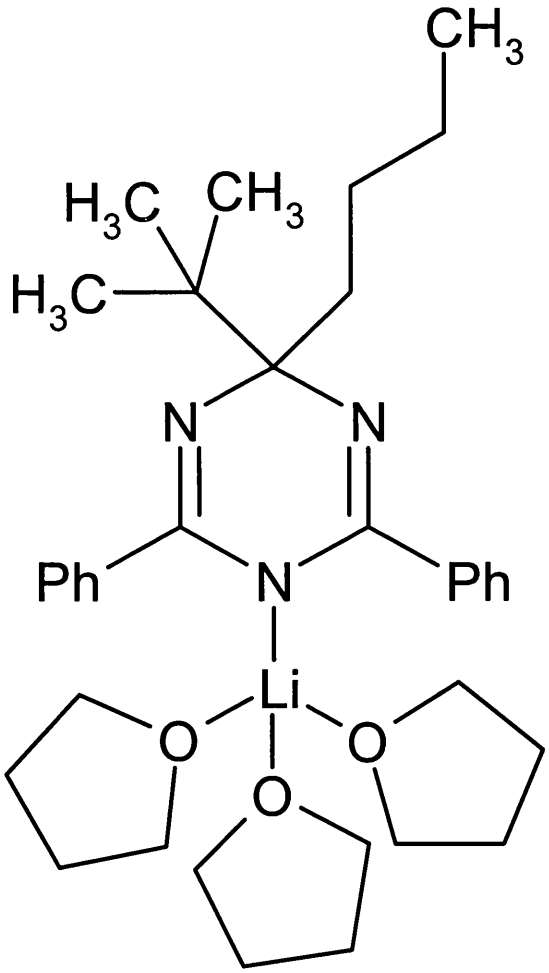
172. Cossi, M. and V. Barone, *Analytical second derivatives of the free energy in solution by polarizable continuum models.* J. Chem. Phys., 1998. **109**(15): p. 6246.

Appendices

Appendix A - CSD reference, formula, compound name, crystallographic information and structure for each of the structures obtained.

LAFBUP10

Reference:	D.R.Armstrong, K.W.Henderson, M.MacGregor, R.E.Mulvey, M.J.Ross, W.Clegg, P.A.O'Neil (1995) *J.Organomet.Chem.*,**486**,79				
Formula:	$C_{35}\ H_{52}\ Li_1\ N_3\ O_3$				
Compound Name:	(4-n-Butyl-4-t-butyl-2,6-diphenyl-1,4-dihydro-1,3,5-triazine)-tris(tetrahydrofuran)-lithium				
Space Group:	P21/n	***Cell:***	***a*** 10.917(2)	***b*** 15.893(3)	***c*** 20.227(4)
Space Group No.:	14 ***(Å,°)***		α 90.00	β 100.30(2)	γ 90.00
R-Factor (%):	9.2	**Temperature(K)**: 295		**Density(g/cm^3)**: 1.096	

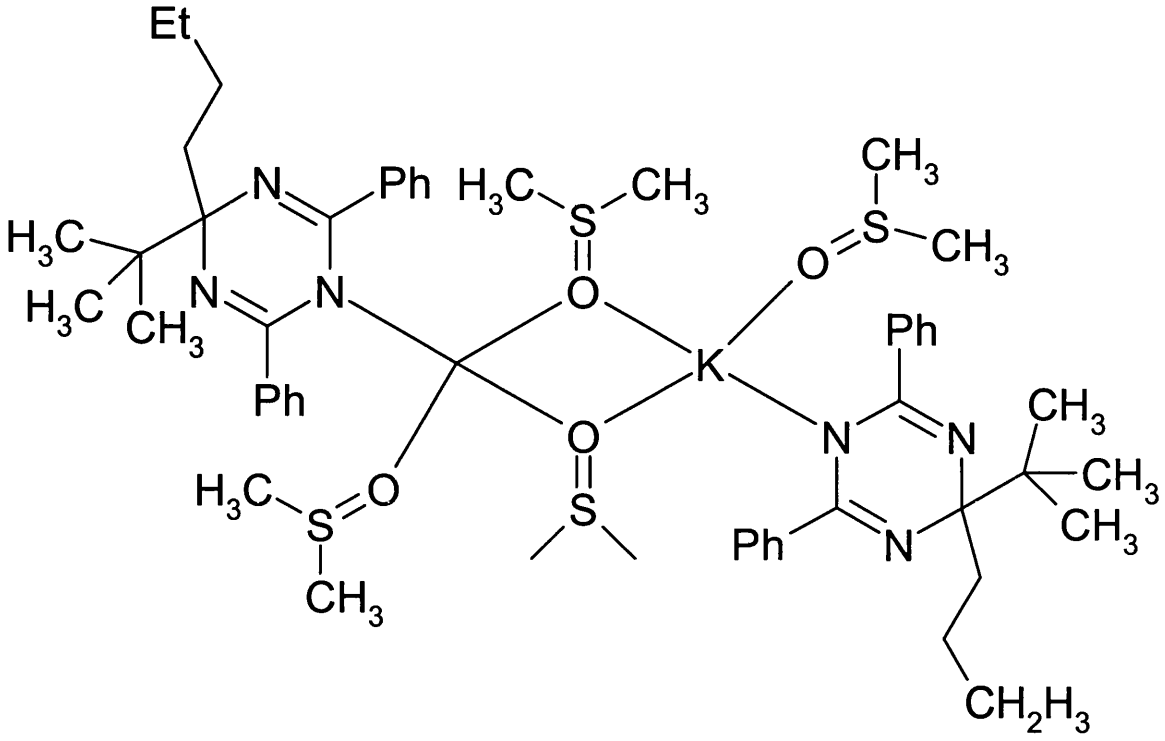


REMGOF

Reference:	W.Clegg, M.R.J.Elsegood, L.Horsburgh, R.E.Mulvey, M.J.Ross (1997) *Chem.Ber.,***130,**621
Formula:	$C_{54}\ H_{80}\ K_2\ N_6\ O_4\ S_4$
Compound Name:	bis((μ2-Dimethylsulfoxide)-(4-n-butyl-4-t-butyl-2,6-diphenyl-1H-1,3,5,-triazinyl)-(dimethylsulfoxide)-potassium)

Space Group:	P21/n	***Cell:***	***a*** 9.147(0)	***b*** 19.138(2)	***c*** 17.178(2)
Space Group No.:	14 ***(Å,°)***		α 90.00	β 93.73(0)	γ 90.00

R-Factor (%):	4.62	***Temperature(K):*** 160	***Density(g/cm^3):*** 1.199



REMGUL

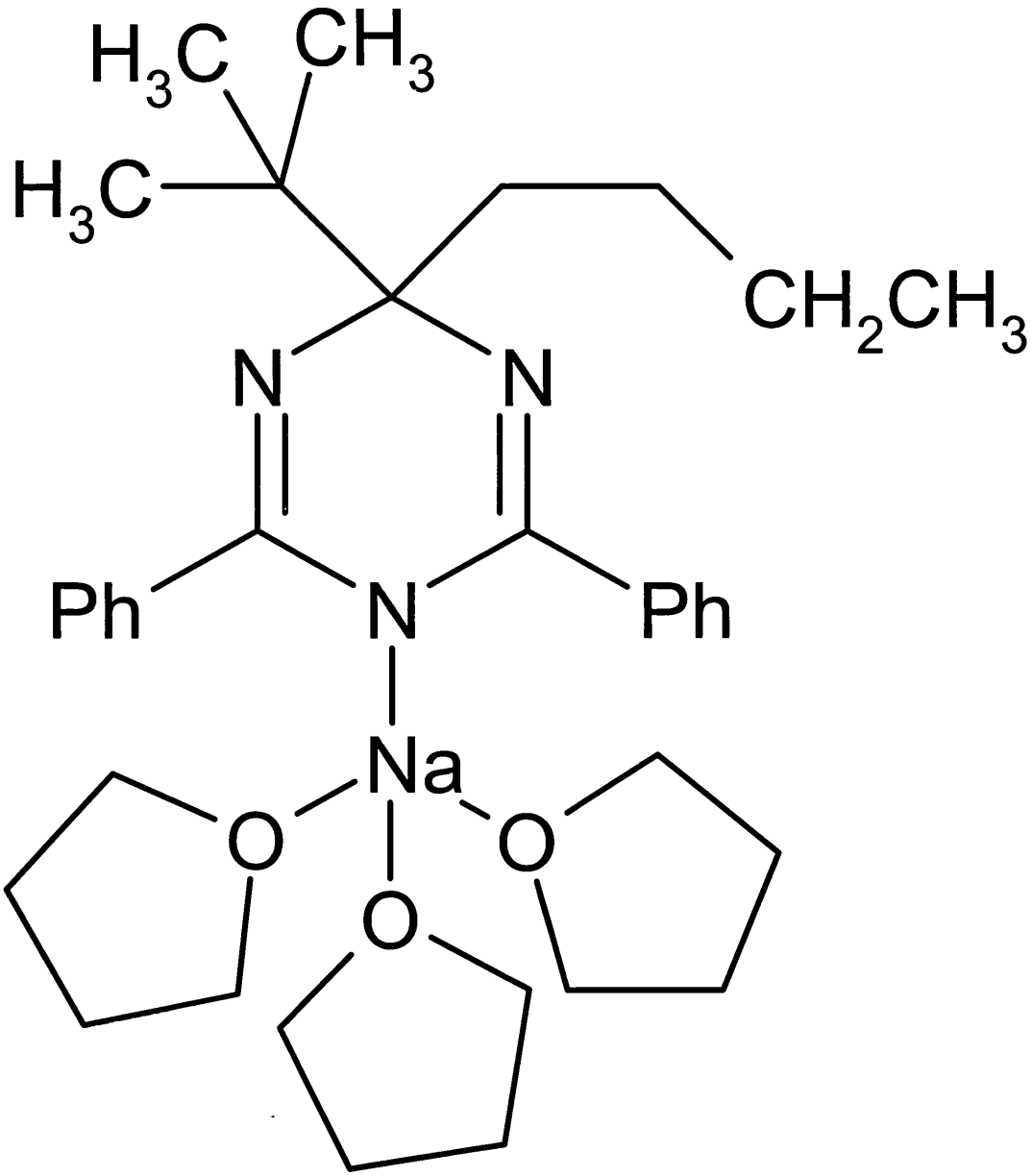
Reference: Clegg, M.R.J.Elsegood, L.Horsburgh, R.E.Mulvey, J.Ross (1997) *Chem.Ber.,***130**,621

Formula: $C_{35}\ H_{52}\ N_3\ Na_1\ O_3$

Compound Name: (4-n-Butyl-4-t-butyl-2,6-diphenyl-1H-1,3,5,-triazine)-tris(tetrahydrofuran)-sodium

Space Group:	P21/n	***Cell:***	***a*** 21.283(4)	***b*** 15.864(3)	***c*** 21.847(4)
Space Group No.:	14 ***(Å,°)***		α 90.00	β 109.55(0)	γ 90.00

R-Factor (%): 9.59 ***Temperature(K):*** 160 ***Density(g/cm3):*** 1.12

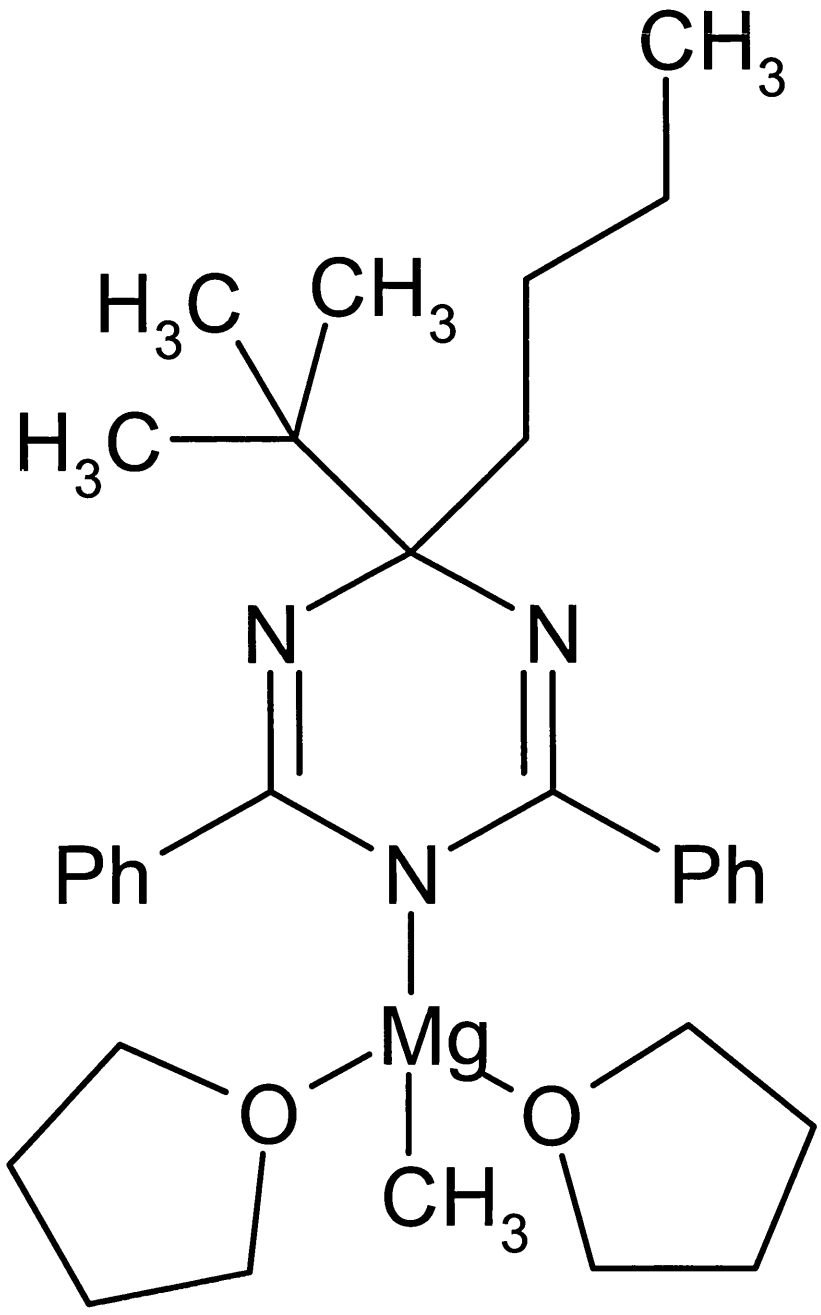


YIZGIX

Reference:	D.R.Armstrong, K.W.Henderson, M.MacGregor, R.E.Mulvey, M.J.Ross, W.Clegg, P.A.O'Neil (1995) *J.Organomet.Chem.*,**486**,79
Formula:	$C_{32}\ H_{47}\ Mg_1\ N_3\ O_2$
Compound Name:	(4-n-Butyl-4-t-butyl-2,6-diphenyl-1,3,5-triazine)-methylbis(tetrahydrofuran)-magnesium

Space Group:	P21/c	***Cell:***	***a*** 12.756(3)	***b*** 15.899(2)	***c*** 16.118(3)
Space Group No.:	14 ***(Å,°)***		α 90.00	β 93.67(0)	γ 90.00

R-Factor (%):	5.7	***Temperature(K):***	295	***Density(g/cm^3):***	1.079



KDMTNB10

Reference:	H.Ueda, N.Sakabe, J.Tanaka (1968) *Bull.Chem.Soc.Jpn.*, **41**,2866				
Formula:	$C_8 H_8 N_3 O_8$ 1-,K_1 1+,2($H_2 O_1$)				
Compound Name:	1,1'-Dimethoxy-2,4,6-trinitrobenzene potassium dihydrate				
Synonym:	Meisenheimer complex				
Space Group:	P-1	***Cell:***	***a*** 9.360(20)	***b*** 10.840(20)	***c*** 7.400(20)
Space Group No.:	2 ***(Å,°)***		α 87.67	β 106.72	γ 102.75(33)
R-Factor (%):	13.0	***Temperature(K):*** 295		***Density(g/cm^3):***	1.654